

# EXACT FORMULAS FOR THE NORMALIZING CONSTANTS OF WISHART DISTRIBUTIONS FOR GRAPHICAL MODELS

BY CAROLINE UHLER<sup>\*</sup>, ALEX LENKOSKI<sup>†</sup>, AND DONALD RICHARDS<sup>§</sup>

*IST Austria<sup>\*</sup>, Norwegian Computing Center<sup>†</sup>, Heidelberg University<sup>§</sup> and  
Penn State University<sup>§</sup>*

Gaussian graphical models have received considerable attention during the past four decades from the statistical and machine learning communities. In Bayesian treatments of this model, the  $G$ -Wishart distribution serves as the conjugate prior for inverse covariance matrices satisfying graphical constraints. While it is straightforward to posit the unnormalized densities, the normalizing constants of these distributions have been known only for graphs that are chordal, or decomposable. Up until now, it was unknown whether the normalizing constant for a general graph could be represented explicitly, and a considerable body of computational literature emerged that attempted to avoid this apparent intractability. We close this question by providing an explicit representation of the  $G$ -Wishart normalizing constant for general graphs.

**1. Introduction.** Let  $G = (V, E)$  be an undirected graph with vertex set  $V = \{1, \dots, p\}$  and edge set  $E$ . Let  $\mathbb{S}^p$  be the set of symmetric  $p \times p$  matrices and  $\mathbb{S}_{\succ 0}^p$  the cone of positive definite matrices in  $\mathbb{S}^p$ . Let

$$(1.1) \quad \mathbb{S}_{\succ 0}^p(G) = \{M = (M_{ij}) \in \mathbb{S}_{\succ 0}^p \mid M_{ij} = 0 \text{ for all } (i, j) \notin E\}$$

denote the cone in  $\mathbb{S}^p$  of positive definite matrices with zeros in all entries not corresponding to edges in the graph. Note that the positivity of all diagonal entries  $M_{ii}$  follows from the positive-definiteness of the matrices  $M$ .

A random vector  $X \in \mathbb{R}^p$  is said to *satisfy the Gaussian graphical model (GGM) with graph  $G$*  if  $X$  has a multivariate normal distribution with mean  $\mu$  and covariance matrix  $\Sigma$ , denoted  $X \sim \mathcal{N}_p(\mu, \Sigma)$ , where  $\Sigma^{-1} \in \mathbb{S}_{\succ 0}^p(G)$ . The inverse covariance matrix  $\Sigma^{-1}$  is called the *concentration matrix* and, throughout this paper, we denote  $\Sigma^{-1}$  by  $K$ .

---

*MSC 2010 subject classifications:* Primary 62H05, 60E05; secondary 62E15

*Keywords and phrases:* Bartlett decomposition; Bipartite graph; Cholesky decomposition; Chordal graph; Directed acyclic graph;  $G$ -Wishart distribution; Gaussian graphical model; Generalized hypergeometric function of matrix argument; Moral graph; Normalizing constant; Wishart distribution.

Statistical inference for the concentration matrix  $K$  constrained to  $\mathbb{S}_{>0}^p(G)$  goes back to Dempster [6], who proposed an algorithm for determining the maximum likelihood estimator [cf., 29]. A Bayesian framework for this problem was introduced by Dawid and Lauritzen [5], who proposed the Hyper-Inverse Wishart (HIW) prior distribution for *chordal* (also known as *decomposable* or *triangulated*) graphs  $G$ .

Chordal graphs enjoy a rich set of properties that led the HIW distribution to be particularly amenable to Bayesian analysis. Indeed, for nearly a decade after their introduction, focus on the Bayesian use of GGMs was placed primarily on chordal graphs [see e.g. 10]. This tractability stems from two causes: the ability to sample directly from HIWs [26], and the ability to calculate their normalizing constants, which are critical quantities when comparing graphs or nesting GGMs in hierarchical structures.

Roverato [27] extended the HIW to general  $G$ . Focusing on  $K$ , Atay-Kayis and Massam [3] further studied this prior distribution. Following Letac and Massam [20], Lenkoski and Dobra [19] termed this distribution the  $G$ -Wishart. For  $D \in \mathbb{S}_{>0}^p(G)$  and  $\delta \in \mathbb{R}$ , the  $G$ -Wishart density has the form

$$f_G(K \mid \delta, D) \propto |K|^{\frac{1}{2}(\delta-2)} \exp(-\frac{1}{2}\text{tr}(KD)) \mathbf{1}_{K \in \mathbb{S}_{>0}^p(G)}.$$

This distribution is conjugate [27] and proper for  $\delta > 1$  [22].

Early work on the  $G$ -Wishart distribution was largely computational in nature [4, 7, 8, 15, 19, 30, 31] and was predicated on two assumptions: first, that a direct sampler was unavailable for this class of models and, second, that the normalizing constant could not be explicitly calculated. Lenkoski [18] developed a direct sampler for  $G$ -Wishart variates, mimicking the algorithm of Dempster [6], thereby resolving the first open question. In this paper, we close the second question by deriving for general graphs  $G$  an explicit formula for the  $G$ -Wishart normalizing constant,

$$C_G(\delta, D) = \int_{\mathbb{S}_{>0}^p(G)} |K|^{\frac{1}{2}(\delta-2)} \exp(-\frac{1}{2}\text{tr}(KD)) \, dK,$$

where  $dK = \prod_{i=1}^p dk_{ii} \cdot \prod_{i < j, (i,j) \in E} dk_{ij}$  denotes the product of differentials corresponding to all distinct non-zero entries in  $K$ .

For notational simplicity, we will consider the integral

$$I_G(\delta, D) = \int_{\mathbb{S}_{>0}^p(G)} |K|^\delta \exp(-\text{tr}(KD)) \, dK,$$

which can be expressed in terms of  $C_G(\delta, D)$  as follows: Denote by  $|E|$  the cardinality of the edge set  $E$ ; by changing variables,  $K \rightarrow 2K$ , one obtains

$$C_G(\delta, D) = 2^{-(\frac{1}{2}m\delta + |E|)} I_G\left(\frac{1}{2}(\delta - 2), D\right).$$

The normalizing constant  $I_G(\delta, D)$  is well-known for *complete graphs*, in which every pair of vertices is connected by an edge. In such cases,

$$(1.2) \quad I_{\text{complete}}(\delta, D) = |D|^{-(\delta + \frac{1}{2}(p+1))} \Gamma_p\left(\delta + \frac{1}{2}(p+1)\right),$$

where

$$(1.3) \quad \Gamma_p(\alpha) = \pi^{p(p-1)/4} \prod_{i=1}^p \Gamma\left(\alpha - \frac{1}{2}(i-1)\right),$$

$\text{Re}(\alpha) > \frac{1}{2}(p-1)$ , is the *multivariate gamma function*. The formula (1.2) has a long history, dating back to Wishart [32], Wishart and Bartlett [33], Ingham [13], Siegel [28, *Hilffsatz* 37], Maass [21], and many derivations of a statistical nature; see Olkin [25] and Giri [9, p. 224].

As noted above,  $I_G(\delta, D)$  is also known for *chordal graphs*. Let  $G$  be chordal, and let  $(T_1, \dots, T_d)$  denote a *perfect sequence* of subsets of  $V$ . Further, let  $S_i = (T_1 \cup \dots \cup T_i) \cap T_{i+1}$ ,  $i = 1, \dots, d-1$ ; then,  $S_1, \dots, S_{d-1}$  are called the *separators* of  $G$ . The subsets  $T_i$  and  $S_i$  are *cliques*, meaning that they are complete graphs, and we denote their cardinalities by  $t_i = |T_i|$  and  $s_i = |S_i|$ . For  $S \subseteq \{1, \dots, m\}$ , let  $D_{SS}$  denote the submatrix of  $D$  corresponding to the rows and columns in  $S$ . Then,

$$(1.4) \quad \begin{aligned} I_G(\delta, D) &= \frac{\prod_{i=1}^d I_{T_i}(\delta, D_{T_i T_i})}{\prod_{j=1}^{d-1} I_{S_j}(\delta, D_{S_j S_j})} \\ &= \frac{\prod_{i=1}^d \left( |D_{T_i T_i}|^{-(\delta + \frac{1}{2}(t_i+1))} \Gamma_{t_i}\left(\delta + \frac{1}{2}(t_i+1)\right) \right)}{\prod_{j=1}^{d-1} \left( |D_{S_j S_j}|^{-(\delta + \frac{1}{2}(s_j+1))} \Gamma_{s_j}\left(\delta + \frac{1}{2}(s_j+1)\right) \right)}. \end{aligned}$$

This result follows because, for a chordal graph  $G$ , the  $G$ -Wishart density function can be factored into a product of density functions [5].

For non-chordal graphs the problem of calculating  $I_G(\delta, D)$  has been open for over 20 years, and much of the computational methodology mentioned above was developed with the objective of either approximating  $I_G(\delta, D)$  or avoiding its calculation. Our result shows that an explicit representation of this quantity is indeed possible. This will enable dramatic increases in methodological applications of  $G$ -Wishart variates for general graphs, as model comparison and hierarchical considerations can now be performed without the need for computationally-intensive approximation methods.

The article proceeds as follows. We first treat the case in which  $D = \mathbb{I}_p$ , the  $p \times p$  identity matrix; in Section 2 we explain the main tools used in this paper by showing how one can compute the integral  $I_G(\delta, \mathbb{I}_p)$  for a specific

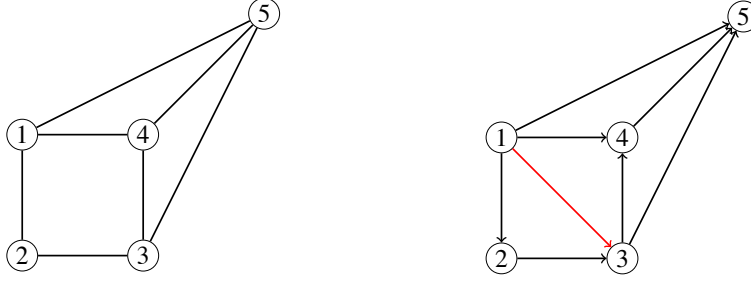


FIG 1. Undirected graph  $G_5$  (left) discussed in Section 2 and its moral DAG  $\tilde{G}_5$  (right) introduced in Section 3.2.

non-chordal graph on five vertices. In Section 3 we derive a general, closed-form, product formula for  $I_G(\delta, \mathbb{I}_p)$ , and we consider in Section 4 the case of general matrices  $D$ .

## 2. Computing $I_G(\delta, \mathbb{I}_m)$ for a non-chordal graph on five vertices.

We consider in this section the non-chordal graph  $G_5$ , shown in Figure 1 (left). The constant  $I_{G_5}(\delta, \mathbb{I}_5)$  is an integral over all positive definite matrices of the form

$$K = \begin{pmatrix} k_{11} & k_{12} & 0 & k_{14} & k_{15} \\ k_{12} & k_{22} & k_{23} & 0 & 0 \\ 0 & k_{23} & k_{33} & k_{34} & k_{35} \\ k_{14} & 0 & k_{34} & k_{44} & k_{45} \\ k_{15} & 0 & k_{35} & k_{45} & k_{55} \end{pmatrix}.$$

PROPOSITION 2.1. *The integral  $I_{G_5}(\delta, \mathbb{I}_5)$  converges absolutely for all  $\delta > -1$ , and*

$$(2.1) \quad I_{G_5}(\delta, \mathbb{I}_5) = \pi^{7/2} \frac{\Gamma(\delta + 1) \Gamma(\delta + \frac{3}{2}) [\Gamma(\delta + 2) \Gamma(\delta + \frac{5}{2})]^2}{\Gamma(\delta + 3)}.$$

PROOF. We write  $K$  as a block matrix

$$K = \begin{pmatrix} K_{AA} & K_{AB} \\ K_{AB}^T & K_{BB} \end{pmatrix},$$

where

$$K_{AA} = \begin{pmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{pmatrix}, \quad K_{AB} = \begin{pmatrix} 0 & k_{14} & k_{15} \\ k_{23} & 0 & 0 \end{pmatrix}, \quad K_{BB} = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{12} & k_{22} & k_{23} \\ k_{13} & k_{23} & k_{33} \end{pmatrix}.$$

According to the determinant formula for block matrices,

$$|K| = |K_{AA}| \cdot |K_{BB} - K_{AB}^T (K_{AA})^{-1} K_{AB}|,$$

and hence

$$\begin{aligned} I_{G_5}(\delta, \mathbb{I}_5) &= \int_{\mathbb{S}_{>0}^m(G_5)} |K|^\delta \exp(-\text{tr}(K)) \, dK \\ &= \int_{\mathbb{S}_{>0}^m(G_5)} |K_{AA}|^\delta |K_{BB} - K_{AB}^T (K_{AA})^{-1} K_{AB}|^\delta \\ &\quad \cdot \exp(-\text{tr}(K_{AA}) - \text{tr}(K_{BB})) \, dK_{AA} \, dK_{AB} \, dK_{BB}. \end{aligned}$$

Shifting  $K_{BB} \rightarrow K_{BB} + K_{AB}^T (K_{AA})^{-1} K_{AB}$  leads to

$$\begin{aligned} I_{G_5}(\delta, \mathbb{I}_5) &= \int_{\mathbb{S}_{>0}^3} |K_{BB}|^\delta \exp(-\text{tr}(K_{BB})) \, dK_{BB} \\ &\quad \cdot \int_{\mathbb{S}_{>0}^2} |K_{AA}|^\delta \exp(-\text{tr}(K_{AA})) \\ &\quad \cdot \int_{\mathbb{R}^3} \exp(-\text{tr}(K_{AB}^T (K_{AA})^{-1} K_{AB})) \, dK_{AB} \, dK_{AA} \\ &= \Gamma_3(\delta + 2) \int_{\mathbb{S}_{>0}^2} |K_{AA}|^\delta \exp(-\text{tr}(K_{AA})) \\ &\quad \cdot \int_{\mathbb{R}^3} \exp(-\text{tr}(K_{AB}^T (K_{AA})^{-1} K_{AB})) \, dK_{AB} \, dK_{AA}, \end{aligned}$$

where we applied Equation (1.2) with  $D = \mathbb{I}_3$  to compute the integral over  $K_{BB}$ .

Denote by  $\text{vec}(K_{AB})$  the vectorized matrix  $K_{AB}$ , written column-by-column. We apply a formula for the Kronecker product of matrices (see Muirhead [24, p. 76]) to obtain

$$\text{tr}(K_{AB}^T (K_{AA})^{-1} K_{AB}) = (\text{vec}(K_{AB}))^T (\mathbb{I}_3 \otimes (K_{AA})^{-1}) \text{vec}(K_{AB}).$$

Let  $\ell_{AB} = (k_{23}, k_{14}, k_{15})^T$  be the column vector containing the non-zero entries of  $\text{vec}(K_{AB})$ , and let  $\Lambda^{-1}$  be the diagonal matrix containing the entries of  $\mathbb{I}_3 \otimes (K_{AA})^{-1}$  corresponding to the components of  $\ell_{AB}$ , i.e.,

$$\begin{aligned} \Lambda^{-1} &= \text{diag}(((K_{AA})^{-1})_{22}, ((K_{AA})^{-1})_{11}, ((K_{AA})^{-1})_{11}) \\ &= |K_{AA}|^{-1} \text{diag}(k_{11}, k_{22}, k_{22}), \end{aligned}$$

where the latter equality follows from the cofactor formula for matrix inverses. Then

$$\text{tr}(K_{AB}^T (K_{AA})^{-1} K_{AB}) = \ell_{AB}^T \Lambda^{-1} \ell_{AB},$$

and hence we obtain the integral over  $K_{AB}$  in the form of a Gaussian integral:

$$\begin{aligned} \int_{\mathbb{R}^3} \exp(-\text{tr}(K_{AB}^T (K_{AA})^{-1} K_{AB})) \, dK_{AB} &= \int_{\mathbb{R}^3} \exp(-\ell_{AB}^T \Lambda^{-1} \ell_{AB}) \, d\ell_{AB} \\ &= \pi^{3/2} |\Lambda|^{1/2} \\ &= \pi^{3/2} |K_{AA}|^{3/2} k_{11}^{-1/2} k_{22}^{-1}. \end{aligned}$$

So we obtain

$$(2.2) \quad I_{G_5}(\delta, \mathbb{I}_5) = \pi^{3/2} \Gamma_3(\delta + 2) \cdot \int_{\mathbb{S}_{>0}^2} |K_{AA}|^{\delta+3/2} k_{11}^{-1/2} k_{22}^{-1} \exp(-\text{tr}(K_{AA})) \, dK_{AA}.$$

To compute this integral, note that  $K_{AA} \in \mathbb{S}_{>0}^2$  if and only if  $k_{11} > 0$ ,  $k_{22} > 0$ , and  $-\sqrt{k_{11}k_{22}} < k_{12} < \sqrt{k_{11}k_{22}}$ . Transforming  $k_{12} \rightarrow \sqrt{k_{11}k_{22}} k_{12}$ , we obtain

$$\begin{aligned} \int_{-\sqrt{k_{11}k_{22}}}^{\sqrt{k_{11}k_{22}}} |K_{AA}|^{\delta+3/2} \, dk_{12} &= \int_{-\sqrt{k_{11}k_{22}}}^{\sqrt{k_{11}k_{22}}} (k_{11}k_{22} - k_{12}^2)^{\delta+3/2} \, dk_{12} \\ &= (k_{11}k_{22})^{\delta+2} \int_{-1}^1 (1 - k_{12}^2)^{\delta+3/2} \, dk_{12} \\ &= (k_{11}k_{22})^{\delta+2} \pi^{1/2} \frac{\Gamma(\delta + \frac{5}{2})}{\Gamma(\delta + 3)}. \end{aligned}$$

Inserting this result into (2.2), we find that the remaining integral with respect to  $k_{11}$  and  $k_{22}$  separates into a product of classical gamma integrals. On evaluating each of those integrals, we obtain

$$I_{G_5}(\delta, \mathbb{I}_5) = \pi^2 \Gamma_3(\delta + 2) \frac{\Gamma(\delta + \frac{5}{2})}{\Gamma(\delta + 3)} \Gamma(\delta + \frac{5}{2}) \Gamma(\delta + 2).$$

Substituting for  $\Gamma_3(\delta + 2)$  from (1.3), we obtain (2.1).  $\square$

In this proof, we relied heavily on the special structure of the graph leading to a block matrix  $K$  with two complete sub-blocks. For example, the shift from  $K_{BB}$  to  $K_{BB} + K_{AB}^T (K_{AA})^{-1} K_{AB}$  requires  $K_{BB}$  to be complete. This proof can be extended to larger graphs with a similar structure. In Section 3 we will use this approach to compute  $I_G(\delta, \mathbb{I})$  for complete bipartite graphs. However, when the graph structure does not lead to a ‘nice’ block structure in  $K$ , it seems difficult to compute  $I_G(\delta, \mathbb{I})$  using this approach.

An alternative approach, which has been suggested and applied previously, is based on the Cholesky decomposition (also known as the Bartlett decomposition): Every positive definite matrix  $K$  can be written uniquely in the form  $K = AA^T$ , where  $A = (a_{ij})$  is an upper-triangular matrix with each  $a_{ii} \in (0, \infty)$  and each off-diagonal  $a_{ij} \in (-\infty, \infty)$ . It is more common to use a lower Cholesky decomposition, i.e.  $K = A^T A$ . However, for reasons which will become clear in Section 3 we will use the upper Cholesky decomposition. Also, for simplifying notation in later sections, we will define the matrix  $A$  to be an upper-triangular matrix with

$$(2.3) \quad A_{ij} = \begin{cases} \sqrt{a_{ii}}, & \text{if } i = j \\ -a_{ij}, & \text{if } i < j. \end{cases}$$

Computing  $AA^T$  with this parametrization yields

$$(AA^T)_{ij} = \begin{cases} a_{ii} + \sum_{l>i} a_{il}^2, & \text{if } i = j \\ -a_{ij}\sqrt{a_{jj}} + \sum_{l>\max(i,j)} a_{il}a_{jl}, & \text{if } i < j. \end{cases}$$

In the following, we show how the integral  $I_{G_5}(\delta, \mathbb{I}_5)$  can be computed using this approach. Let

$$A = \begin{pmatrix} \sqrt{a_{11}} & -a_{12} & -a_{13} & -a_{14} & -a_{15} \\ 0 & \sqrt{a_{22}} & -a_{23} & -a_{24} & -a_{25} \\ 0 & 0 & \sqrt{a_{33}} & -a_{34} & -a_{35} \\ 0 & 0 & 0 & \sqrt{a_{44}} & -a_{45} \\ 0 & 0 & 0 & 0 & \sqrt{a_{55}} \end{pmatrix},$$

and for  $\mathbb{S}_{>0}^p(G)$ , we equate each entry  $k_{ij}$  of  $K$  to the corresponding entry of  $AA^T$ . This results in a set of 15 equations for the entries of  $K$  in terms of the entries of  $A$ . In particular, we find that

$$\begin{aligned} 0 = k_{13} &= -a_{13}\sqrt{a_{33}} + a_{14}a_{34} + a_{15}a_{35}, \\ 0 = k_{24} &= -a_{24}\sqrt{a_{44}} + a_{25}a_{45}, \\ 0 = k_{25} &= -a_{25}\sqrt{a_{55}}. \end{aligned}$$

Since  $k_{25} = 0$  and  $a_{55} > 0$ , then  $a_{25} \equiv 0$ . Also, since  $k_{24} = 0$  and  $a_{44} > 0$ , then  $a_{24} \equiv 0$ . Finally, since  $k_{13} = 0$ , then

$$(2.4) \quad a_{13} = \frac{a_{14}a_{34} + a_{15}a_{35}}{\sqrt{a_{33}}}.$$

Therefore, the system of equations for the non-zero entries of  $K$  reduces to

$$k_{ii} = a_{ii} + \sum_{j>i} a_{ij}^2 \quad \text{and} \quad k_{ij} = -a_{ij}\sqrt{a_{jj}} + \sum_{l>\max(i,j)} a_{il}a_{jl},$$

where we set  $a_{25} = a_{35} = 0$  and replace  $a_{13}$  by  $(a_{14}a_{34} + a_{15}a_{35})/\sqrt{a_{33}}$ .

Next we calculate the Jacobian of the transformation from  $K$  to  $A$ . We list the  $a_{ij}$ 's column-wise, but omitting  $a_{13}$ ,  $a_{24}$ , and  $a_{25}$ , which correspond to the zero's in  $K$ . So,  $a_{11}, a_{12}, a_{22}, a_{23}, a_{33}, a_{14}, a_{34}, a_{44}, a_{15}, a_{35}, a_{45}, a_{55}$  represent columns 1,  $\dots$ , 12, respectively, of  $J$ , the  $12 \times 12$  Jacobian matrix. Similarly,  $k_{11}, k_{12}, k_{22}, k_{23}, k_{33}, k_{14}, k_{34}, k_{44}, k_{15}, k_{35}, k_{45}, k_{55}$  represent rows 1,  $\dots$ , 12, respectively, of  $J$ . In calculating  $J$  using the partial derivative of each  $k_{ij}$  with respect to each  $a_{lm}$ , one can easily see that  $J$  is upper-triangular, so we obtain  $\det(J) = \det(\text{diag}(J)) = a_{22}^{1/2} a_{33}^{1/2} a_{44} a_{55}^{3/2}$ .

We now have all the ingredients necessary to calculate

$$(2.5) \quad I_{G_5}(\delta, \mathbb{I}_5) = \int_{K \in \mathbb{S}_{>0}^5(G_5)} |K|^\delta \exp(-\text{tr}(K)) \, dK$$

through the change of variables,  $K = AA^T$ . Then

$$\begin{aligned} \det(K) &= a_{11}a_{22}a_{33}a_{44}a_{55}, \\ \text{tr}(K) &= a_{11} + a_{12}^2 + \left( \frac{a_{14}a_{34} + a_{15}a_{35}}{\sqrt{a_{33}}} \right)^2 + a_{14}^2 + a_{15}^2 + a_{22} + a_{23}^2 \\ &\quad + a_{33} + a_{34}^2 + a_{35}^2 + a_{44} + a_{45}^2 + a_{55}, \end{aligned}$$

and so the integral (2.5) equals

$$\begin{aligned} &\int_A a_{11}^\delta a_{22}^{\delta+1/2} a_{33}^{\delta+1/2} a_{44}^{\delta+1} a_{55}^{\delta+3/2} \\ &\quad \times \exp \left[ - \left( a_{11} + a_{12}^2 + \left( \frac{a_{14}a_{34} + a_{15}a_{35}}{\sqrt{a_{33}}} \right)^2 + a_{14}^2 + a_{15}^2 \right. \right. \\ &\quad \left. \left. + a_{22} + a_{23}^2 + a_{33} + a_{34}^2 + a_{35}^2 + a_{44} + a_{45}^2 + a_{55} \right) \right] \, dA, \end{aligned}$$

where  $a_{ii} > 0$ ;  $a_{ij} \in \mathbb{R}$ ,  $i < j$ ; and  $dA$  denotes the product of all differentials.

The integrals with respect to  $a_{ii}$ ,  $i = 1, 2, 4, 5$  each are gamma integrals of the form

$$\int_0^\infty a^\delta \exp(-a) \, da = \Gamma(\delta + 1), \quad \delta > -1.$$



Now we integrate with respect to all variables corresponding to edges, except those variables involved in Equation (2.4) for  $a_{13}$ . This yields a Gaussian integral for each variable  $a_{12}$ ,  $a_{23}$ , and  $a_{45}$ :

$$\int_{-\infty}^{\infty} \exp(-a^2) da = \sqrt{\pi}.$$

Finally, we have an integral over all variables involved in Equation (2.4):

$$\begin{aligned} \int \cdots \int_{\substack{a_{33} > 0 \\ a_{14}, a_{15}, a_{34}, a_{35} \in \mathbb{R}}} \exp \left[ - \left( \left( \frac{a_{14}a_{34} + a_{15}a_{35}}{\sqrt{a_{33}}} \right)^2 + a_{14}^2 + a_{15}^2 + a_{34}^2 + a_{35}^2 \right) \right] \\ \times a_{33}^{\delta+1/2} \exp(-a_{33}) da_{14} da_{15} da_{34} da_{35} da_{33}. \end{aligned}$$

Let us view  $a_{14}$  and  $a_{15}$  as independent, identically distributed (i.i.d.)  $\mathcal{N}(0, 1)$  random variables. Then, for fixed  $a_{34}, a_{35}, a_{33}$ , the random variable

$$\frac{a_{14}a_{34} + a_{15}a_{35}}{\sqrt{a_{33}}} \sim \mathcal{N} \left( 0, \frac{a_{34}^2 + a_{35}^2}{a_{33}} \right),$$

and hence,

$$\left( \frac{a_{14}a_{34} + a_{15}a_{35}}{\sqrt{a_{33}}} \right)^2 \sim \frac{a_{34}^2 + a_{35}^2}{a_{33}} Y,$$

where  $Y \sim \chi_1^2$ . Therefore,

$$\begin{aligned} \iint_{a_{14}, a_{15} \in \mathbb{R}} \exp \left[ - \left( \left( \frac{a_{14}a_{34} + a_{15}a_{35}}{\sqrt{a_{33}}} \right)^2 + a_{14}^2 + a_{15}^2 \right) \right] da_{14} da_{15} \\ = \pi \mathbb{E} \left[ \exp \left( - \frac{a_{34}^2 + a_{35}^2}{2a_{33}} Y \right) \right] = \pi \left( 1 + \frac{a_{34}^2 + a_{35}^2}{a_{33}} \right)^{-1/2}. \end{aligned}$$

So, the remaining integrals are

$$\pi \iiint_{\substack{a_{33} > 0 \\ a_{34}, a_{35} \in \mathbb{R}}} a_{33}^{\delta+1/2} \left( 1 + \frac{a_{34}^2 + a_{35}^2}{a_{33}} \right)^{-1/2} \exp \left[ - (a_{33} + a_{34}^2 + a_{35}^2) \right] da_{34} da_{35} da_{33}.$$

Here, we change variables, replacing  $(a_{34}, a_{35})$  by  $(\sqrt{a_{33}}a_{34}, \sqrt{a_{33}}a_{35})$ ; then the latter integral becomes

$$\pi \iiint_{\substack{a_{33} > 0 \\ a_{34}, a_{35} \in \mathbb{R}}} a_{33}^{\delta+3/2} (1 + a_{34}^2 + a_{35}^2)^{-1/2} \exp \left[ -a_{33}(1 + a_{34}^2 + a_{35}^2) \right] da_{33} da_{34} da_{35},$$

and integrating over  $a_{33}$  using the classical gamma integral yields

$$\pi \Gamma\left(\delta + \frac{5}{2}\right) \iint_{a_{34}, a_{35} \in \mathbb{R}} (1 + a_{34}^2 + a_{35}^2)^{-(\delta+3)} da_{34} da_{35}$$

which, by applying (2.6) below, equals

$$\pi^2 \Gamma\left(\delta + \frac{5}{2}\right) \frac{\Gamma(\delta + 2)}{\Gamma(\delta + 3)}.$$

Collecting all terms, we obtain the same result as in Proposition 2.1.

For the last step in the above computation we used a well-known result: For  $\rho > k/2$ ,

$$(2.6) \quad \int_{\mathbb{R}^k} (1 + v^T v)^{-\rho} dv = \pi^{k/2} \frac{\Gamma(\rho - \frac{1}{2}k)}{\Gamma(\rho)}.$$

This result can be proved by changing variables to polar coordinates and evaluating the resulting integral by elementary methods; alternatively, the integral can be deduced from the density function of the multivariate  $t$ -distribution (see [1, p. 55]).

A clear advantage of the approach which uses a Cholesky decomposition of  $K$  is that it allows us to easily integrate out all variables that are not involved in equations occurring due to zeros in  $K$ , e.g., Equation (2.4) for  $a_{13}$  in the above computation. However, for cases in which there are many equations of this type, or the variables in these equations overlap, the computation of  $I_G(\delta, \mathbb{I}_p)$  using this approach will generally be difficult. For both approaches, the ease of computation depends crucially on the ordering of the columns and rows, or equivalently, on the labeling of the vertices in the graph. We will see in Section 3 that one can directly deduce from the labeled graph the number and form of the equations that complicate the computations (such as Equation (2.4) in the case of  $G_5$ ). The labeling or ordering of the variables that yields a minimal number of such equations is called a *minimal fill-in ordering* and is explained in the following section.

**3. Computing  $I_G(\delta, \mathbb{I}_p)$  for general non-chordal graphs.** In this section, we compute  $I_G(\delta, \mathbb{I}_p)$  for general non-chordal graphs. We start in Section 3.1 with the class of complete bipartite graphs. Because of the special structure of these graphs, we can use a similar approach as in the proof of Proposition 2.1 in the previous section. However, it seems that this approach does not easily lead to a formula for general non-chordal graphs. In

Section 3.2 we introduce directed Gaussian graphical models and show how these models relate to the Cholesky factor approach to computing  $I_G(\delta, \mathbb{I}_p)$ . This will lead, ultimately, to a procedure for computing the normalizing constant for any graph  $G$ .

3.1. *Bipartite graphs.* A complete bipartite graph on  $m + n$  vertices, denoted by  $H_{m,n}$ , is an undirected graph whose vertices can be divided into two disjoint sets,  $U = \{1, \dots, m\}$  and  $V = \{m + 1, \dots, m + n\}$ , such that every vertex in  $U$  is connected to all vertices in  $V$ , but there are no edges within  $U$  or within  $V$ . For the graph  $H_{m,n}$ , the corresponding matrix  $K$  is a block matrix,

$$K = \begin{pmatrix} K_{AA} & K_{AB} \\ K_{AB}^T & K_{BB} \end{pmatrix},$$

where  $K_{AA}$  and  $K_{BB}$  are diagonal matrices of sizes  $m \times m$  and  $n \times n$ , respectively, and  $K_{AB}$  is complete.

PROPOSITION 3.1. *The integral  $I_{H_{m,n}}(\delta, \mathbb{I}_{m+n})$  converges absolutely for all  $\delta > -1$ , and*

$$(3.1) \quad I_{H_{m,n}}(\delta, \mathbb{I}_{m+n}) = [\Gamma(\delta + \tfrac{1}{2}n + 1)]^m [\Gamma(\delta + \tfrac{1}{2}m + 1)]^n \\ \times \frac{\Gamma_{m+n}(\delta + \tfrac{1}{2}(m + n + 1))}{\Gamma_m(\delta + \tfrac{1}{2}(m + n + 1)) \Gamma_n(\delta + \tfrac{1}{2}(m + n + 1))}.$$

PROOF. Applying the determinant formula for block matrices, we obtain

$$I_{H_{m,n}}(\delta, \mathbb{I}_{m+n}) = \int_{\mathbb{S}_{>0}^{m+n}(G)} |K|^\delta \exp(-\text{tr}(K)) dK \\ = \int_{\mathbb{S}_{>0}^{m+n}(G)} |K_{AA}|^\delta |K_{BB} - K_{AB}^T(K_{AA})^{-1}K_{AB}|^\delta \\ \cdot \exp(-\text{tr}(K_{AA}) - \text{tr}(K_{BB})) dK_{AA} dK_{AB} dK_{BB}.$$

Since  $K_{AB}$  is complete, we can make a change of variables and replace  $K_{AB}$  by  $K_{AA}^{1/2}K_{AB}K_{BB}^{1/2}$ ; the corresponding Jacobian is  $|K_{AA}|^{n/2}|K_{BB}|^{m/2}$ . Since

$$|K_{BB} - K_{BB}^{1/2}K_{AB}^T K_{AB} K_{BB}^{1/2}| = |K_{BB}| \cdot |\mathbb{I}_n - K_{AB}^T K_{AB}|,$$

we obtain

$$I_{H_{m,n}}(\delta, \mathbb{I}_{m+n}) = \int_{\mathbb{S}_{>0}^{m+n}(G)} |K_{AA}|^{\delta + \frac{1}{2}n} |K_{BB}|^{\delta + \frac{1}{2}m} |\mathbb{I}_n - K_{AB}^T K_{AB}|^\delta \\ \cdot \exp(-\text{tr}(K_{AA}) - \text{tr}(K_{BB})) dK_{AA} dK_{AB} dK_{BB},$$

where the range of integration is such that each diagonal entry of  $K_{AA}$  and  $K_{BB}$  is positive,  $K_{AB}$  is complete, and  $\mathbb{I}_n - K_{AB}^T K_{AB}$  is positive definite.

Integrating over each diagonal entry of  $K_{AA}$  and  $K_{BB}$ , we obtain

$$I_{H_{m,n}}(\delta, \mathbb{I}_{m+n}) = [\Gamma(\delta + \tfrac{1}{2}n + 1)]^m [\Gamma(\delta + \tfrac{1}{2}m + 1)]^n \\ \times \int_{K_{AB}} |\mathbb{I}_n - K_{AB}^T K_{AB}|^\delta dK_{AB}.$$

Finally, since  $K_{AB}$  is complete, we deduce the latter integral from (4.4).  $\square$

In this computation, we used the special structure of the graph to decompose the inverse covariance matrix  $K$  into a special block matrix. However, it seems difficult to extend this approach to examples for which  $K_{AB}$  is not complete. Therefore, we focus in the sequel on the approach based on the Cholesky factorization of  $K$ , as described in Section 2.

**3.2. Directed Gaussian graphical models.** Let  $\mathcal{G} = (V, \mathcal{E})$  be a directed acyclic graph (DAG) consisting of vertices  $V = \{1, \dots, p\}$  and directed edges  $\mathcal{E}$ . We assume, without loss of generality, that the vertices in  $\mathcal{G}$  are *topologically ordered*, meaning that  $i < j$  for all  $(i, j) \in \mathcal{E}$ . We associate to  $\mathcal{G}$  a strictly upper-triangular matrix  $B$  of edge weights. So  $B = (b_{ij})$  with  $b_{ij} \neq 0$  if and only if  $(i, j) \in \mathcal{E}$ . Then a *directed Gaussian graphical model* on  $\mathcal{G}$  for a random variable  $X \in \mathbb{R}^p$  is defined by  $X \sim \mathcal{N}_p(0, \Sigma)$  with  $\Sigma = [(I - B)D(I - B)^T]^{-1}$ , where  $D$  is a diagonal matrix.

To simplify notation, let  $a_{ii} = d_{ii}$  and  $a_{ij} = -b_{ij}\sqrt{d_{jj}}$ , and let  $A = (A_{ij})$  with  $A_{ii} = \sqrt{a_{ii}}$  and  $A_{ij} = -a_{ij}$  for all  $i \neq j$ . This is the same parametrization as introduced in (2.3). Then  $\Sigma^{-1} = AA^T$ , and  $a_{ij} \neq 0$  for  $i \neq j$  if and only if  $(i, j) \in \mathcal{E}$ . Note that  $AA^T$  is the upper Cholesky decomposition of  $\Sigma^{-1}$ . Such a decomposition exists for any positive definite matrix and is unique.

We will associate to a DAG,  $\mathcal{G} = (V, \mathcal{E})$ , and its corresponding directed Gaussian graphical model two undirected graphs. We denote by  $\mathcal{G}^s = (V, \mathcal{E}^s)$  the *skeleton* of  $\mathcal{G}$  obtained by replacing all directed edges in  $\mathcal{G}$  by undirected edges. We denote by  $\mathcal{G}^m = (V, \mathcal{E}^m)$  the *moral graph* of  $\mathcal{G}$ , which reflects the conditional independencies in  $\mathcal{N}_p(0, \Sigma)$ , i.e.,

$$(i, j) \notin \mathcal{E}^m \quad \text{if and only if} \quad X_i \perp\!\!\!\perp X_j \mid X_{V \setminus \{i, j\}}.$$

Since  $\Sigma^{-1}$  also encodes the conditional independence relations of the form  $X_i \perp\!\!\!\perp X_j \mid X_{V \setminus \{i, j\}}$ , this is equivalent to the criterion,

$$(i, j) \notin \mathcal{E}^m \quad \text{if and only if} \quad (\Sigma^{-1})_{ij} = 0.$$

So, the moral graph  $\mathcal{G}^m$  reflects the zero pattern of  $\Sigma^{-1}$ .

The moral graph of  $\mathcal{G}$  can also be defined graph-theoretically: It is formed by connecting all nodes  $i, j \in V$  that have a common child in  $\mathcal{G}$ , i.e., for which there exists a node  $k \in V \setminus \{i, j\}$  such that  $(i, k), (j, k) \in \mathcal{E}$ , and then making all edges in the graph undirected. The name stems from the fact that the moral graph is obtained by ‘marrying’ the parents. For a review of basic graph-theoretic concepts see e.g. Lauritzen [17, ch. 2].

The moral graph is an important concept for our application. Let  $G = (V, E)$  be an undirected graph with  $V = \{1, \dots, p\}$ , for which we would like to compute  $I_G(\delta, \mathbb{I}_p)$ . Let  $G_0 = (V, E_0)$  with  $G_0 = G$ . Given a labeling of the vertices  $V$  we can associate a DAG  $\mathcal{G}_0 = (V, \mathcal{E}_0)$  to  $G_0$  by orienting the edges in  $E_0$  according to the topological ordering, i.e., for all  $(i, j) \in E_0$  let  $(i, j) \in \mathcal{E}_0$  if  $i < j$ . Note that the skeleton of  $\mathcal{G}_0$  is the original undirected graph  $G_0$ . Let  $G_1 = (V, E_1)$  be the moral graph of  $\mathcal{G}_0$ , i.e.,  $G_1 = \mathcal{G}_0^m$ , and let  $\mathcal{G}_1 = (V, \mathcal{E}_1)$  be the corresponding DAG obtained by orienting the edges in  $E_1$  according to the ordering of the vertices  $V$ . So  $\mathcal{G}_0$  is a subgraph of  $\mathcal{G}_1$ . We repeat this procedure until  $\mathcal{G}_{q+1} = \mathcal{G}_q$ . This results in a sequence of DAGs,

$$\mathcal{G}_0 \subsetneq \mathcal{G}_1 \subsetneq \dots \subsetneq \mathcal{G}_q.$$

In the following, we denote by  $\mathcal{G} = (V, \mathcal{E})$  the DAG associated to  $G = (V, E)$  obtained by orienting the edges in  $E$  according to the ordering of the vertices  $V$ . We denote by  $\bar{\mathcal{G}} = (V, \bar{\mathcal{E}})$  the DAG associated to  $G = (V, E)$  obtained by repeatedly marrying parents in  $\mathcal{G}$ , i.e.  $\bar{\mathcal{G}} = \mathcal{G}_q$ . We call  $\bar{\mathcal{G}}$  the *moral DAG* of  $G$ .

In the following lemma, we prove that  $\bar{\mathcal{G}}^s$ , the skeleton of  $\bar{\mathcal{G}}$ , is a chordal graph with  $G \subset \bar{\mathcal{G}}^s$ . So  $\bar{\mathcal{G}}^s$  is a *chordal cover* of  $G$ . A chordal cover in general is not unique. However,  $\bar{\mathcal{G}}^s$  is the unique chordal cover obtained by repeatedly marrying parents according to the vertex labeling  $V$ . We call this chordal cover the *moral chordal graph* of  $G$  and denote it by  $\bar{G} = (V, \bar{E})$ .

**LEMMA 3.2.** *Let  $G = (V, E)$  be an undirected graph and let  $\bar{\mathcal{G}} = (V, \bar{\mathcal{E}})$  be the moral DAG of  $G$  obtained by repeatedly marrying parents. Then  $\bar{G}$ , the skeleton of  $\bar{\mathcal{G}}$ , is a chordal graph.*

**PROOF.** We assume that  $\bar{G}$  is not chordal. So there exists a cycle of size at least 4 in  $\bar{G}$  that does not contain any chord. Since  $\bar{\mathcal{G}}$  is a DAG, the corresponding cycle in  $\bar{\mathcal{G}}$  must contain three vertices  $i, j, l$  with  $(i, l), (j, l) \in \bar{\mathcal{E}}$ . Since the cycle does not contain any chords, this means that the parents  $i, j$  are not connected in  $\bar{\mathcal{G}}$ . This is a contradiction to  $\bar{\mathcal{G}}$  being the moral DAG of  $G$ .  $\square$

We now show how we can deduce from the undirected graph  $G = (V, E)$  the normalizing constant  $I_G(\delta, \mathbb{I}_p)$  as an integral in terms of the Cholesky factor  $A$ . In the following, we use the standard graph-theoretic notation  $\text{indeg}(i)$  for the *indegree* of node  $i$ , representing the number of edges “arriving at” (or “pointing to”) node  $i$  in a DAG  $\mathcal{G}$ .

**THEOREM 3.3.** *Let  $G = (V, E)$  be an undirected graph with vertices  $V = \{1, \dots, p\}$ . Let  $\mathcal{G} = (V, \mathcal{E})$  be the DAG associated to  $G = (V, E)$  obtained by orienting the edges in  $E$  according to the ordering of the vertices in  $V$ . Let  $\bar{\mathcal{G}} = (V, \bar{\mathcal{E}})$  denote the moral DAG of  $G$  and  $\bar{G} = (V, \bar{E})$  its skeleton, the moral chordal graph of  $G$ . Let  $A$  be an upper-triangular matrix of size  $p \times p$  with diagonal entries  $A_{ii} = \sqrt{a_{ii}}$  and off-diagonal entries  $A_{ij} = -a_{ij}$  for all  $i < j$ . Then*

$$I_G(\delta, \mathbb{I}_p) = \int_{A_*} \left( \prod_{i=1}^p a_{ii}^{\delta + \frac{1}{2} \text{indeg}(i)} \right) \exp \left[ - \sum_{i=1}^p a_{ii} - \sum_{(i,j) \in \bar{\mathcal{E}}} a_{ij}^2 \right] dA_*,$$

where  $A_* = \{a_{ij} : i = j \text{ or } (i, j) \in \mathcal{E}\}$ , the range of  $a_{ii}$  is  $(0, \infty)$ , the range of  $a_{ij}$  for  $(i, j) \in \mathcal{E}$  is  $(-\infty, \infty)$ , and for  $a_{ij} \notin A_*$ ,

$$a_{ij} = \begin{cases} 0, & \text{if } (i, j) \notin \bar{\mathcal{E}} \\ \frac{1}{\sqrt{a_{jj}}} \sum_{\substack{l \in V \\ (i,l), (j,l) \in \bar{\mathcal{E}}}} a_{il}a_{jl}, & \text{if } (i, j) \in \bar{\mathcal{E}} \setminus \mathcal{E} \end{cases}$$

**PROOF.** Let  $K \in \mathbb{S}_{>0}^p(G)$ . Since  $G \subset \bar{G}$ , then  $K \in \mathbb{S}_{>0}^p(\bar{G})$  and we can view  $K$  as an inverse covariance matrix of a directed Gaussian graphical model on  $\bar{\mathcal{G}}$ . Because the Cholesky decomposition is unique,  $A$  is a weighted adjacency matrix of  $\bar{\mathcal{G}}$  and hence  $a_{ij} = 0$  for all  $(i, j) \notin \bar{\mathcal{E}}$ .

Let  $(i, j)$  be an edge that is present in the moral chordal graph  $\bar{G}$  but not in  $G$ . We can assume that  $i < j$ . Hence  $(i, j) \in \bar{\mathcal{E}} \setminus \mathcal{E}$  and therefore

$$0 = K_{ij} = (AA^T)_{ij} = -a_{ij}\sqrt{a_{jj}} + \sum_{l > \max(i,j)} a_{il}a_{jl}.$$

Thus, for each edge  $(i, j) \in \bar{\mathcal{E}} \setminus \mathcal{E}$ , we obtain an equation,

$$a_{ij} = \frac{1}{\sqrt{a_{jj}}} \sum_{\substack{l \in V \\ (i,l), (j,l) \in \bar{\mathcal{E}}}} a_{il}a_{jl}.$$

To complete the proof, we need to compute the Jacobian  $J$  of the transformation from  $K$  to  $A$ . We list the  $a_{ij}$ 's column-wise, meaning that  $a_{ij}$

precedes  $a_{lm}$  if  $j < m$  or if  $j = m$  and  $i < l$ , omitting  $a_{ij}$  for  $(i, j) \notin \mathcal{E}$ , corresponding to the zeros in  $K$ . We list the  $k_{ij}$ 's in the same ordering. Let the  $a_{ij}$ 's correspond to the columns of the Jacobian, while the  $k_{ij}$ 's correspond to the rows. In order to form  $J$ , we calculate the partial derivative of each  $k_{ij}$  with respect to each  $a_{lm}$ . Since  $K = AA^T$  and  $A$  is upper-triangular then  $J$  also is upper-triangular; therefore,  $|J| = |\text{diag}(J)|$ . Since

$$k_{ii} = a_{ii} + \sum_{(i,j) \in \bar{\mathcal{E}}} a_{ij}^2 \quad \text{and} \quad k_{ij} = -a_{ij}\sqrt{a_{jj}} + \sum_{\substack{l \in V \\ (i,l), (j,l) \in \bar{\mathcal{E}}}} a_{il}a_{jl},$$

for all  $(i, j) \in \mathcal{E}$ , then

$$|J| = \prod_{i=1}^p a_{ii}^{\text{indeg}(i)/2}.$$

Putting everything together completes the proof.  $\square$

EXAMPLE 3.4. We revisit the example  $G_5$  discussed in Section 2. The moral DAG of  $G_5$  is denoted by  $\bar{\mathcal{G}}_5$  and depicted in Figure 1 (right). Since the edges  $(2, 4)$  and  $(2, 5)$  are missing in  $\bar{\mathcal{G}}_5$ , we immediately deduce that  $a_{24} = a_{25} = 0$ . In this example, only one edge needed to be added in the process of marrying parents, namely the edge  $(1, 3)$ . This results in one equation for  $a_{13}$ , which can be deduced from the *colliders* over the additional edge, i.e., nodes  $l \in V$  with  $(1, l), (3, l) \in \bar{\mathcal{G}}$ , and results in

$$a_{13} = \frac{1}{\sqrt{a_{33}}}(a_{14}a_{34} + a_{15}a_{35}).$$

Finally, the Jacobian can be deduced from the indegrees of the nodes in  $\mathcal{G}_5$ , which corresponds to the moral DAG  $\bar{\mathcal{G}}_5$  after omitting the red edge. So the determinant of the Jacobian is

$$a_{11}^{0/2} a_{22}^{1/2} a_{33}^{1/2} a_{44}^{2/2} a_{55}^{3/2}.$$

When computing the integral  $I_{G_5}(\delta, \mathbb{I}_5)$  in Section 2, we saw that the equations corresponding to the additional edges  $(i, j) \in \bar{\mathcal{E}} \setminus \mathcal{E}$  complicate the computations significantly. So it is desirable to choose an ordering which requires us to add as few edges as possible. In fact, one can find an ordering of the vertices such that  $\bar{\mathcal{G}} = \mathcal{G}$  if and only if  $G$  is chordal. This is proven in the following Lemma.

LEMMA 3.5. *Let  $G = (V, E)$  be an undirected graph. Then  $G$  is chordal if and only if there exists an ordering of the nodes such that  $\bar{\mathcal{G}} = \mathcal{G}$ .*

PROOF. We first assume that  $G$  is not chordal. Since for every ordering of the vertices, the skeleton of  $\bar{\mathcal{G}}$  is chordal,  $\bar{\mathcal{G}} \neq \mathcal{G}$ . Next, we assume that  $G$  is chordal. Let  $(T_1, \dots, T_d)$  denote a perfect sequence of subsets of  $V$ . We order the vertices such that  $i < j$  if  $i \in T_l, j \in T_m$  and  $l < m$ . Within each subset  $T_k$ , the ordering does not matter. With this ordering all parents in  $\mathcal{G}$  are already married and hence  $\bar{\mathcal{G}} = \mathcal{G}$ .  $\square$

Therefore, no “complicated” equations need to be added if and only if  $G$  is chordal. In the following Theorem, we show how one can directly derive the normalizing constant  $I_G(\delta, \mathbb{I}_p)$  from the graph  $G$  when  $G$  is chordal. One could also prove this result by using Equation (1.4).

PROPOSITION 3.6. *Let  $G = (V, E)$  be a chordal graph, where the vertices  $V = \{1, \dots, p\}$  are labelled according to a perfect ordering. Then*

$$I_G(\delta, \mathbb{I}_p) = \pi^{|E|/2} \prod_{i=1}^p \Gamma(\delta + \tfrac{1}{2} \text{indeg}(i) + 1).$$

where  $\text{indeg}(i)$  denotes the indegree of node  $i$  in the DAG  $\mathcal{G}$ .

PROOF. It follows from Lemma 3.5 that since  $G$  is chordal and the vertices are labelled according to a perfect ordering,  $\bar{\mathcal{G}} = \mathcal{G}$ . Hence, by Theorem 3.3,

$$I_G(\delta, \mathbb{I}_p) = \int_{A_*} \left( \prod_{i=1}^p a_{ii}^{\delta + \frac{1}{2} \text{indeg}(i)} \right) \exp \left[ - \sum_{i=1}^p a_{ii} - \sum_{(i,j) \in \mathcal{E}} a_{ij}^2 \right] dA_*,$$

where  $A_* = (a_{ij} : i = j \text{ or } (i, j) \in \mathcal{E})$ . Integrating variable by variable, we obtain

$$\begin{aligned} I_G(\delta, \mathbb{I}_p) &= \prod_{i=1}^p \int_0^\infty a_{ii}^{\delta + \frac{1}{2} \text{indeg}(i)} \exp(-a_{ii}) da_{ii} \cdot \prod_{(i,j) \in \mathcal{E}} \int_{-\infty}^\infty \exp(-a_{ij}^2) da_{ij} \\ &= \pi^{|E|/2} \prod_{i=1}^p \Gamma(\delta + \tfrac{1}{2} \text{indeg}(i) + 1). \end{aligned}$$

The proof now is complete.  $\square$

This result shows that it is straightforward to integrate over edges and nodes that are not involved in the equations for the additional edges  $(i, j) \in \bar{\mathcal{E}} \setminus \mathcal{E}$  needed to make a graph chordal. So, given a non-chordal graph  $G$ , it is helpful to find an ordering such that  $|\bar{\mathcal{E}} \setminus \mathcal{E}|$  is minimized. This ordering is



given by a perfect ordering of a minimal chordal cover of  $G$ , where minimality is with respect to the number of edges that need to be added in order to make  $G$  chordal. Using Proposition 3.6, we can compute the normalizing constant corresponding to a minimal chordal cover of  $G$ . The question remains, how can one compute the normalizing constant of  $G$  from the normalizing constant of a minimal chordal cover of  $G$ ? In the following theorem, we show how one can compute the normalizing constant of a graph  $G$  from the normalizing constant of the graph resulting from adding one additional edge  $e$  to  $G$ .

**THEOREM 3.7.** *Let  $G = (V, E)$  be an undirected graph with vertices  $V = \{1, \dots, p\}$ . Let  $G^e = (V, E^e)$  denote the graph  $G$  with one additional edge  $e$ , i.e.,  $E^e = E \cup \{e\}$ . Let  $d$  denote the number of triangles formed by the edge  $e$  and two other edges in  $G^e$ . Then*

$$I_G(\delta, \mathbb{I}_p) = \pi^{-1/2} \frac{\Gamma(\delta + \frac{1}{2}(d+2))}{\Gamma(\delta + \frac{1}{2}(d+3))} I_{G^e}(\delta, \mathbb{I}_p).$$

**PROOF.** We begin by defining an ordering of the vertices in such a way that one can easily integrate out the variables corresponding to the end points of  $e$  and the variable corresponding to  $e$  itself.

Let one of the end points of  $e$  be labelled as ‘1’, the other end point as ‘ $d+2$ ’ and label the  $d$  vertices involved in triangles over the edge  $e$  by  $2, \dots, d+1$ . Label all remaining vertices by  $d+3, \dots, p$ . Let  $\bar{\mathcal{G}}^e$  denote the moral DAG to  $G^e$  with edge set  $\bar{\mathcal{E}}^e$ . Then the chosen ordering of the vertices guarantees that  $\bar{\mathcal{E}}^e = \bar{\mathcal{E}} \cup \{e\}$ , and  $e \notin \bar{\mathcal{E}}$ .

Also, since all vertices  $2, \dots, d+1$  are connected to vertex  $d+2$ , no added edge in  $\bar{\mathcal{E}} \setminus \mathcal{E}$  points to vertex  $d+2$  and hence  $a_{d+2, d+2}$  does not appear in any equation for the edges in  $\bar{\mathcal{E}} \setminus \mathcal{E}$ . Similar arguments hold for vertex 1, since due to the ordering there can be no edge pointing to node 1.

Let  $A$  denote the Cholesky factor of  $G$  and  $A^e$  the Cholesky factor of  $G^e$ . Then

$$A_{ij} = \begin{cases} A_{ij}^e & \text{for all } (i, j) \neq (1, d+2) \\ 0 & \text{if } (i, j) = (1, d+2) \end{cases}$$

Let  $\text{indeg}$  denote the indegree with respect to the DAG  $\mathcal{G}$  and  $\text{indeg}^e$  the indegree with respect to the DAG  $\mathcal{G}^e$ . Let  $A_* = ((a_{ii})_{i \notin \{1, d+2\}}, (a_{ij})_{(i, j) \in \mathcal{E}})$ . Note that

$$(3.2) \quad \text{indeg}^e(1) = 0 = \text{indeg}(1), \quad \text{indeg}^e(d+2) = d+1 = \text{indeg}(d+2) + 1.$$

Then by Theorem 3.3,

$$\begin{aligned}
I_{G^e}(\delta, \mathbb{I}_p) &= \int \left( \prod_{i=1}^p a_{ii}^{\delta + \frac{1}{2} \text{indeg}^e(i)} \exp(-a_{ii}) \right) \exp \left[ - \sum_{(i,j) \in \bar{\mathcal{E}}^e} a_{ij}^2 \right] \\
&\quad da_{11} da_{d+2,d+2} da_{1,d+2} dA_* \\
&= \int_{-\infty}^{\infty} \exp(-a_{1,d+2}^2) da_{1,d+2} \\
&\quad \cdot \int_0^{\infty} a_{11}^{\delta + \frac{1}{2} \text{indeg}^e(1)} \exp(-a_{11}) da_{11} \\
&\quad \cdot \int_0^{\infty} a_{d+2,d+2}^{\delta + \frac{1}{2} \text{indeg}^e(d+2)} \exp(-a_{d+2,d+2}) da_{d+2,d+2} \\
&\quad \cdot \int_{A_*} \left[ \prod_{i \notin \{1,d+2\}} a_{ii}^{\delta + \frac{1}{2} \text{indeg}^e(i)} \exp(-a_{ii}) \right] \exp \left[ - \sum_{(i,j) \in \bar{\mathcal{E}}} a_{ij}^2 \right] dA_*.
\end{aligned}$$

The integral with respect to  $a_{1,d+2}$  is a Gaussian integral, with value  $\sqrt{\pi}$ . Also, by (3.2),

$$\int_0^{\infty} a_{11}^{\delta + \frac{1}{2} \text{indeg}^e(1)} \exp(-a_{11}) da_{11} = \int_0^{\infty} a_{11}^{\delta + \frac{1}{2} \text{indeg}(1)} \exp(-a_{11}) da_{11}.$$

Again by (3.2), we have

$$\begin{aligned}
&\int_0^{\infty} a_{d+2,d+2}^{\delta + \frac{1}{2} \text{indeg}^e(d+2)} \exp(-a_{d+2,d+2}) da_{d+2,d+2} \\
&= \frac{\Gamma(\delta + \frac{1}{2}(d+1) + 1)}{\Gamma(\delta + \frac{1}{2}d + 1)} \int_0^{\infty} a_{d+2,d+2}^{\delta + \frac{1}{2} \text{indeg}(d+2)} \exp(-a_{d+2,d+2}) da_{d+2,d+2}.
\end{aligned}$$

Finally, since  $\text{indeg}^e(i) = \text{indeg}(i)$  for all  $i \notin \{1, d+2\}$ , we obtain

$$\begin{aligned}
I_{G^e}(\delta, \mathbb{I}_p) &= \sqrt{\pi} \frac{\Gamma(\delta + \frac{1}{2}(d+1) + 1)}{\Gamma(\delta + \frac{1}{2}d + 1)} \int_0^{\infty} a_{11}^{\delta + \text{indeg}(1)/2} \exp(-a_{11}) da_{11} \\
&\quad \cdot \int_0^{\infty} a_{d+2,d+2}^{\delta + \frac{1}{2} \text{indeg}(d+2)} \exp(-a_{d+2,d+2}) da_{d+2,d+2} \\
&\quad \cdot \int_{A_*} \left( \prod_{i \notin \{1,d+2\}} a_{ii}^{\delta + \frac{1}{2} \text{indeg}(i)} \exp(-a_{ii}) \right) \exp \left[ - \sum_{(i,j) \in \bar{\mathcal{E}}} a_{ij}^2 \right] dA_* \\
&= \sqrt{\pi} \frac{\Gamma(\delta + \frac{1}{2}(d+3))}{\Gamma(\delta + \frac{1}{2}(d+2))} I_G(\delta, \mathbb{I}_p).
\end{aligned}$$

The proof now is complete.  $\square$

By applying this theorem we can compute the normalizing constant for any graph  $G$ . First, we find a chordal cover of  $G$  and compute its normalizing constant by applying Proposition 3.6. It is desirable to start with a *minimal* chordal cover where, as before, “minimality” refers to the number of added edges. Then we delete, one at a time, the added edges that are in the chordal cover of  $G$ , but not in  $G$ , and compute the corresponding normalizing constant by applying Theorem 3.7. This finally results in the normalizing constant for  $G$ . In the following, we show how this procedure works for the graph  $G_5$ .

EXAMPLE 3.8. We revisit the example  $G_5$  discussed in Section 2. The skeleton of the graph shown in Figure 1 (right) is a chordal cover of  $G_5$  and the given vertex labeling is a perfect labeling. By applying Proposition 3.6, we deduce the normalizing constant for the graph  $G_5$  with the additional edge  $e = (1, 3)$ :

$$I_{G_5^e}(\delta, \mathbb{I}_p) = \pi^4 \Gamma(\delta + 1) \Gamma(\delta + \tfrac{3}{2}) [\Gamma(\delta + 2)]^2 \Gamma(\delta + \tfrac{5}{2}).$$

Since the number of triangles over the red edge  $(1, 3)$  is  $d = 3$ , we find by Theorem 3.7 that

$$\begin{aligned} I_{G_5}(\delta, \mathbb{I}_p) &= \pi^{-1/2} \frac{\Gamma(\delta + \tfrac{3}{2} + 1)}{\Gamma(\delta + \tfrac{4}{2} + 1)} I_{G_5^e}(\delta, \mathbb{I}_p) \\ &= \pi^{7/2} \frac{\Gamma(\delta + \tfrac{5}{2})}{\Gamma(\delta + 3)} \Gamma(\delta + 1) \Gamma(\delta + \tfrac{3}{2}) [\Gamma(\delta + 2)]^2 \Gamma(\delta + \tfrac{5}{2}). \end{aligned}$$

**4. Computing  $I_G(\delta, D)$  for general non-chordal graphs.** In this section we extend the results in Section 3 to general  $D \in \mathbb{S}_{>0}^p$  and give a similar characterization as in Theorem 3.7, showing how the normalizing constant changes when removing an edge from a graph. This results in a procedure for computing  $I_G(\delta, D)$  for any graph  $G$ .

4.1. *Some results on a generalized hypergeometric function of matrix argument.* We list in this subsection some results, involving a generalized hypergeometric function of matrix argument, that we will apply repeatedly in this section.

For  $a \in \mathbb{C}$  and  $k \in \{0, 1, 2, \dots\}$ , we denote the *rising factorial* by

$$(a)_k = \frac{\Gamma(a + k)}{\Gamma(a)} = a(a + 1)(a + 2) \cdots (a + k - 1).$$

For  $t \in \mathbb{C}$  and  $\rho \notin \{0, -1, -2, \dots\}$  the *classical generalized hypergeometric function*,  ${}_0F_1(\rho, t)$ , may be defined by the series expansion,

$$(4.1) \quad {}_0F_1(\rho; t) = \sum_{l=0}^{\infty} \frac{t^l}{l! (\rho)_l}.$$

We refer to Andrews, et al. [2] for many other properties of this function.

The *generalized hypergeometric function of matrix argument*,  ${}_0F_1(\rho; Y)$ ,  $Y \in \mathbb{S}_{>0}^p$ , is defined by the Laplace transform,

$$\frac{1}{\Gamma_p(\rho)} \int_{\mathbb{S}_{>0}^p} |Y|^{\rho - \frac{1}{2}(p+1)} \exp(-\text{tr}(YD)) {}_0F_1(\rho; Y) dY = |D|^{-\rho} \exp(\text{tr}(D^{-1})),$$

valid for  $\text{Re}(\rho) > \frac{1}{2}(p-1)$  and  $D \in \mathbb{S}_{>0}^p$ . Herz [12] provided an extensive theory of the analytic properties of the function  ${}_0F_1$ . In particular,  ${}_0F_1(\rho; Y)$  is simultaneously analytic in  $\rho$  for  $\text{Re}(\rho) > \frac{1}{2}(p-1)$  and entire in  $Y$ ; so, as a function of  $Y$ , its domain of definition extends to the set  $\mathbb{S}^p$  and to the set of complex symmetric matrices. Other properties of the function  ${}_0F_1$ , such as zonal polynomial expansions which generalize (4.1), are given by James [14], Muirhead [24], and Gross and Richards [11].

Herz [12, p. 497] proved that the function  ${}_0F_1(\rho; Y)$  depends only on the eigenvalues of  $Y$ , and moreover that if  $\text{Re}(\rho) > \frac{1}{2}(p-1)$ ,  $D \in \mathbb{S}_{>0}^p$ , and  $C \in \mathbb{S}^p$ , then there holds the Laplace transform formula,

$$(4.2) \quad \int_{\mathbb{S}_{>0}^p} |Y|^{\rho - \frac{1}{2}(p+1)} \exp(-\text{tr}(YD)) {}_0F_1(\rho; YC) dY \\ = \Gamma_p(\rho) |D|^{-\rho} \exp(\text{tr}(D^{-1}C)),$$

where, by convention,  ${}_0F_1(\rho; YC)$  is an abbreviation for  ${}_0F_1(\rho; Y^{1/2}CY^{1/2})$  and  $Y^{1/2} \in \mathbb{S}_{>0}^p$  is the unique square-root of  $Y$ . Setting  $C = 0$  (the zero matrix) in (4.2) we deduce from the uniqueness of the Laplace transform and (1.2) that  ${}_0F_1(\rho; 0) = 1$ .

We will apply repeatedly a generalization of the Poisson integral to matrix spaces (see [12, pp. 495–496] and [14, Equation (151)]): If  $A$  is a  $k \times p$  matrix such that  $k \leq p$ , and  $\text{Re}(\rho) > \frac{1}{2}(k+p-1)$ , then

$$(4.3) \quad \int_{0 < XX^T < \mathbb{I}_k} |\mathbb{I}_k - XX^T|^{\rho - \frac{1}{2}(k+p+1)} \exp(\text{tr}(AX^T)) dX \\ = \frac{\pi^{kp/2} \Gamma_k(\rho - \frac{1}{2}p)}{\Gamma_k(\rho)} {}_0F_1\left(\rho; \frac{1}{4}AA^T\right),$$

where the region of integration is the set of all  $k \times p$  matrices  $X$  such that  $XX^T \in \mathbb{S}_{>0}^k$  and  $I - XX^T \in \mathbb{S}_{>0}^k$ . In particular, on setting  $A = 0$  we obtain

$$(4.4) \quad \int_{0 < XX^T < \mathbb{I}_k} |\mathbb{I}_k - XX^T|^{\rho - \frac{1}{2}(k+p+1)} dX = \frac{\pi^{kp/2} \Gamma_k(\rho - \frac{1}{2}p)}{\Gamma_k(\rho)},$$

a result which was used in Proposition 3.1.

For the case in which  $Y$  is a  $2 \times 2$  matrix, Muirhead [23] proved that

$$(4.5) \quad {}_0F_1(\rho; Y) = \sum_{q=0}^{\infty} \frac{1}{q! (\rho)_{2q} (\rho - \frac{1}{2})_q} |Y|^q {}_0F_1(\rho + 2q; \text{tr}(Y)),$$

where the  ${}_0F_1$  functions on the right-hand side are the classical generalized hypergeometric functions given in (4.1). In the special case in which  $Y$  is of rank 1, it follows from Herz [12, p. 497], or directly from (4.5), that

$$(4.6) \quad {}_0F_1(\rho; Y) = {}_0F_1(\rho; \text{tr}(Y)).$$

4.2. *The normalizing constant for non-chordal graphs.* We want to calculate

$$I_G(\delta, D) = \int_{\mathbb{S}_{>0}^p(G)} |K|^\delta \exp(-\text{tr}(KD)) dK,$$

the normalizing constant for  $G$ , a general non-chordal graph. By making the change of variables  $K \rightarrow \text{diag}(D)^{-1/2} K \text{diag}(D)^{-1/2}$  we can assume, without loss of generality, that  $D$  has ones on the diagonal and therefore is a correlation matrix; this assumption will be maintained explicitly for the remainder of the paper.

In the sequel, we will encounter a  $2 \times m$  matrix  $C = (C_{ij})$ , and then we use the notation  $|C_{\{1,2\},\{i,j\}}|$  for the minor corresponding to rows 1 and 2 and to columns  $i$  and  $j$ , where  $i, j \in \{1, \dots, m\}$ . We will need  $L = (L_{ij})$ , a  $2 \times m$  matrix of non-negative integers such that  $\sum_{i=1}^2 \sum_{j=1}^m L_{ij} = l$ , and we adopt the notation

$$\binom{l}{L} = \frac{l!}{\prod_{i=1}^2 \prod_{j=1}^m L_{ij}!}, \quad L_{i+} = \sum_{j=1}^m L_{ij}, \quad \text{and} \quad L_{+j} = \sum_{i=1}^2 L_{ij}.$$

We will also have  $Q = (Q_{ij})_{1 \leq i < j \leq m}$ , a vector of non-negative integers such that  $\sum_{1 \leq i < j \leq m} Q_{ij} = q$ , and we set

$$\binom{q}{Q} = \frac{q!}{\prod_{1 \leq i < j \leq m} Q_{ij}!}, \quad Q_{i+} = \sum_{j=i+1}^m Q_{ij}, \quad \text{and} \quad Q_{+j} = \sum_{i=1}^{j-1} Q_{ij}.$$

In the following result, we obtain the normalizing constant for  $H_{2,m}$ , a complete bipartite graph on  $2 + m$  vertices.

PROPOSITION 4.1. *The integral  $I_{H_{2,m}}(\delta, D)$  converges absolutely for all  $\delta > -1$  and  $D \in \mathbb{S}_{>0}^{2+m}$ . Let  $C = (C_{ij})$  denote the  $2 \times m$  submatrix of  $D$  corresponding to the edges in  $G$ ; then  $I_{H_{2,m}}(\delta, D)$  is given by*

$$\begin{aligned}
I_{H_{2,m}}(\delta, \mathbb{I}_{m+2}) &= \sum_{q=0}^{\infty} \frac{(\delta + \frac{1}{2}(m+2))_q [(\delta+2)_q]^m}{q! (\delta + \frac{1}{2}(m+3))_{2q}} \sum_{l=0}^{\infty} \frac{1}{l! (\delta + 2q + \frac{1}{2}(m+3))_l} \\
&\cdot \sum_L \binom{l}{L} \left( \prod_{i=1}^2 \prod_{j=1}^m C_{ij}^{L_{ij}} \right) \left( \prod_{i=1}^2 (\delta + q + \frac{1}{2}(m+2))_{L_{i+}} \right) \left( \prod_{j=1}^m (\delta + 2)_{L_{+j}} \right) \\
&\cdot \sum_Q \binom{q}{Q} \left( \prod_{1 \leq i < j \leq m} |C_{\{1,2\},\{i,j\}}|^{2Q_{ij}} \right) \left( \prod_{j=1}^m (\delta + L_{+j} + 2)_{Q_{j+} + Q_{+j}} \right),
\end{aligned}$$

with

$$(4.7) \quad I_{H_{2,m}}(\delta, \mathbb{I}_{m+2}) = \frac{\pi^m \Gamma_2(\delta + \frac{3}{2})}{\Gamma_2(\delta + \frac{1}{2}(m+3))} [\Gamma(\delta+2)]^m \Gamma(\delta + \frac{1}{2}(m+2))^2.$$

PROOF. We order the vertices such that

$$K = \begin{pmatrix} K_{AA} & K_{AB} \\ K_{BA} & K_{BB} \end{pmatrix},$$

where  $K_{AA} = \text{diag}(\kappa_1, \kappa_2)$ ,  $K_{BB} = \text{diag}(k_1, \dots, k_m)$ , and  $K_{AB}$  is complete. We partition  $D$  in a similar way

$$D = \begin{pmatrix} D_{AA} & D_{AB} \\ D_{BA} & D_{BB} \end{pmatrix},$$

where  $\text{diag}(D) = (1, \dots, 1)$  and  $D_{AB} = C$ . By applying the determinant formula for block matrices and making a change of variables to replace  $K_{AB}$  by  $K_{AA}^{1/2} K_{AB} K_{BB}^{1/2}$ , we obtain similarly as in the proof of Proposition 3.1:

$$\begin{aligned}
I_{H_{2,m}}(\delta, D) &= \int_{\mathbb{S}_{>0}^{2+m}(G)} |K|^\delta \exp(-\text{tr}(KD)) dK \\
&= \int_{\mathbb{S}_{>0}^{2+m}(G)} |K_{AA}|^{\delta + \frac{1}{2}m} |K_{BB}|^{\delta+1} |\mathbb{I}_m - K_{AB}^T K_{AB}|^\delta \\
&\quad \cdot \exp(-\text{tr}(K_{AA}) - \text{tr}(K_{BB})) \\
&\quad \cdot \exp \left[ -2\text{tr} \left( K_{AA}^{1/2} K_{AB} K_{BB}^{1/2} C^T \right) \right] dK_{AA} dK_{AB} dK_{BB}.
\end{aligned}$$

Applying (4.3) to integrate over  $K_{AB}$ , we obtain

$$\begin{aligned} I_G(\delta, D) &= \frac{\pi^m \Gamma_2(\delta + \frac{3}{2})}{\Gamma_2(\delta + \frac{1}{2}(m+3))} \\ &\quad \cdot \int |K_{AA}|^{\delta + \frac{1}{2}m} |K_{BB}|^{\delta+1} \exp(-\text{tr}(K_{AA}) - \text{tr}(K_{BB})) \\ &\quad \cdot {}_0F_1(\delta + \frac{1}{2}(m+3); K_{AA}CK_{BB}C^T) dK_{AA} dK_{BB}. \end{aligned}$$

Applying (4.5) and (4.1) we get

$$\begin{aligned} &{}_0F_1(\delta + \frac{1}{2}(m+3); K_{AA}CK_{BB}C^T) \\ &= \sum_{q=0}^{\infty} \frac{1}{q! (\delta + \frac{1}{2}(m+3))_{2q} (\delta + \frac{1}{2}(m+2))_q} |K_{AA}CK_{BB}C^T|^q \\ &\quad \cdot {}_0F_1(\delta + 2q + \frac{1}{2}(m+3); \text{tr}(K_{AA}CK_{BB}C^T)) \\ &= \sum_{q=0}^{\infty} \frac{1}{q! (\delta + \frac{1}{2}(m+3))_{2q} (\delta + \frac{1}{2}(m+2))_q} |K_{AA}CK_{BB}C^T|^q \\ &\quad \cdot \sum_{l=0}^{\infty} \frac{1}{l! (\delta + 2q + \frac{1}{2}(m+3))_l} (\text{tr}(K_{AA}CK_{BB}C^T))^l. \end{aligned}$$

By the Binet-Cauchy formula (see Karlin [16, p. 1]),

$$\begin{aligned} |K_{AA}CK_{BB}C^T| &= |K_{AA}| \cdot |CK_{BB}C^T| \\ &= |K_{AA}| \sum_{1 \leq i < j \leq m} k_i k_j |C_{\{1,2\},\{i,j\}}|^2. \end{aligned}$$

Hence, by the Multinomial Theorem,

$$\begin{aligned} &|K_{AA}CK_{BB}C^T|^q \\ &= |K_{AA}|^q \sum_Q \binom{q}{Q} \prod_{1 \leq i < j \leq m} (k_i k_j |C_{\{1,2\},\{i,j\}}|^2)^{Q_{ij}} \\ &= |K_{AA}|^q \sum_Q \binom{q}{Q} \left( \prod_{i=1}^m k_i^{Q_{i+} + Q_{+i}} \right) \left( \prod_{1 \leq i < j \leq m} |C_{\{1,2\},\{i,j\}}|^{2Q_{ij}} \right), \end{aligned}$$

where  $Q = (Q_{ij})_{1 \leq i < j \leq m}$  is a vector of non-negative integers, as defined earlier. Also,

$$\text{tr}(K_{AA}CK_{BB}C^T) = \sum_{i=1}^2 \sum_{j=1}^m \kappa_i k_j C_{ij},$$

and hence, by the Multinomial Theorem,

$$\begin{aligned}
(\text{tr}(K_{AA}CK_{BB}C^T))^l &= \left( \sum_{i=1}^2 \sum_{j=1}^m \kappa_i k_j C_{ij} \right)^l \\
&= \sum_L \binom{l}{L} \prod_{i=1}^2 \prod_{j=1}^m (\kappa_i k_j C_{ij})^{L_{ij}} \\
&= \sum_L \binom{l}{L} \left( \prod_{i=1}^2 (\kappa_i)^{L_{i+}} \right) \left( \prod_{j=1}^m k_j^{L_{+j}} \right) \left( \prod_{i=1}^2 \prod_{j=1}^m (C_{ij})^{L_{ij}} \right),
\end{aligned}$$

where  $L = (L_{ij})$  is a  $2 \times m$  matrix of non-negative integers, as defined earlier. Therefore,

$$\begin{aligned}
I_{H_{2,m}}(\delta, D) &= \frac{\pi^m \Gamma_2(\delta + \frac{3}{2})}{\Gamma_2(\delta + \frac{1}{2}(m+3))} \sum_{q=0}^{\infty} \frac{1}{q! (\delta + \frac{1}{2}(m+3))_{2q} (\delta + \frac{1}{2}(m+2))_q} \\
&\quad \cdot \sum_{l=0}^{\infty} \frac{1}{l! (\delta + 2q + \frac{1}{2}(m+3))_l} \\
&\quad \cdot \sum_L \binom{l}{L} \left( \prod_{i=1}^2 \prod_{j=1}^m C_{ij}^{L_{ij}} \right) \left( \prod_{i=1}^2 \int_0^{\infty} \kappa_i^{\delta+q+L_{i+}+\frac{1}{2}m} e^{-\kappa_i} d\kappa_i \right) \\
&\quad \cdot \sum_Q \binom{q}{Q} \left( \prod_{1 \leq i < j \leq m} |C_{\{1,2\},\{i,j\}}|^{2Q_{ij}} \right) \\
&\quad \cdot \left( \prod_{j=1}^m \int_0^{\infty} k_j^{\delta+Q_{j+}+Q_{+j}+L_{+j}+1} e^{-k_j} dk_j \right).
\end{aligned}$$

By evaluating each gamma integral and simplifying the outcomes, we obtain

$$\begin{aligned}
I_{H_{2,m}}(\delta, D) &= \frac{\pi^m \Gamma_2(\delta + \frac{3}{2})}{\Gamma_2(\delta + \frac{1}{2}(m+3))} [\Gamma(\delta + 2)]^m [\Gamma(\delta + \frac{1}{2}(m+2))]^2 \\
&\quad \cdot \sum_{q=0}^{\infty} \frac{(\delta + \frac{1}{2}(m+2))_q ((\delta + 2)_q)^m}{q! (\delta + \frac{1}{2}(m+3))_{2q}} \sum_{l=0}^{\infty} \frac{1}{l! (\delta + 2q + \frac{1}{2}(m+3))_l} \\
&\quad \cdot \sum_L \binom{l}{L} \left( \prod_{i=1}^2 \prod_{j=1}^m (C_{ij})^{L_{ij}} \right) \left( \prod_{i=1}^2 (\delta + q + \frac{1}{2}(m+2))_{L_{i+}} \right) \left( \prod_{j=1}^m (\delta + 2)_{L_{+j}} \right) \\
&\quad \cdot \sum_Q \binom{q}{Q} \left( \prod_{1 \leq i < j \leq m} |C_{\{1,2\},\{i,j\}}|^{2Q_{ij}} \right) \left( \prod_{j=1}^m (\delta + L_{+j} + 2)_{Q_{j+}+Q_{+j}} \right).
\end{aligned}$$



Finally, the value of  $I_{H_{2,m}}(\delta, \mathbb{I}_{m+2})$  is obtained by applying Theorem 3.1 or Theorem 3.7, so the proof now is complete.  $\square$

Note that if we set  $D = \mathbb{I}_{m+2}$  in the proof of Proposition 4.1 then  $|C_{\{1,2\},\{i,j\}}| = C_{ij} = 0$ . Hence, in the infinite series, the only non-zero terms are those for which  $l = q = 0$ , so the series reduces identically to 1.

The special structure of  $K$  was crucial for the proof of Proposition 4.1. For general graphs we will need a combination of the approach used in this proof, where we represent  $K$  as a block matrix, and the approach used earlier, of representing  $K$  by its upper Cholesky decomposition. We now describe the normalizing constant as an integral over the variables in the Cholesky decomposition of  $K$ . This result is a generalization of Theorem 3.3.

**THEOREM 4.2.** *Let  $G = (V, E)$  be an undirected graph with vertices  $V = \{1, \dots, p\}$ . Let  $\mathcal{G} = (V, \mathcal{E})$  be the DAG associated to  $G = (V, E)$  obtained by orienting the edges in  $E$  according to the ordering of the vertices in  $V$ . Let  $\bar{\mathcal{G}} = (V, \bar{\mathcal{E}})$  denote the moral DAG of  $G$ . Let  $A$  be an upper-triangular matrix of size  $p \times p$  with diagonal entries  $A_{ii} = \sqrt{a_{ii}}$  and off-diagonal entries  $A_{ij} = -a_{ij}$  for all  $i < j$ . Then*

$$I_G(\delta, D) = \int_{A_*} \left( \prod_{i=1}^p a_{ii}^{\delta + \frac{1}{2} \text{indeg}(i)} \right) \exp \left[ - \sum_{i=1}^p \left( a_{ii} + \sum_{j: (i,j) \in \bar{\mathcal{E}}} a_{ij}^2 \right) \right] \\ \cdot \exp \left[ - 2 \sum_{(i,j) \in \mathcal{E}} d_{ij} \left( -a_{ij} \sqrt{a_{jj}} + \sum_{k: (i,k), (j,k) \in \bar{\mathcal{E}}} a_{ik} a_{jk} \right) \right] dA_*,$$

where  $D \in \mathbb{S}_{>0}^p$  is a correlation matrix,  $A_* = \{a_{ij} : i = j \text{ or } (i, j) \in \mathcal{E}\}$ , the range of  $a_{ii}$  is  $(0, \infty)$ , the range of  $a_{ij}$  for  $(i, j) \in \mathcal{E}$  is  $(-\infty, \infty)$ ,  $\text{indeg}(i)$  denotes the indegree of node  $i$  in  $\mathcal{G}$ , and for  $a_{ij} \notin A_*$

$$a_{ij} = \begin{cases} 0, & \text{if } (i, j) \notin \bar{\mathcal{E}}, \\ \frac{1}{\sqrt{a_{jj}}} \sum_{\substack{k \in V \\ (i,k), (j,k) \in \bar{\mathcal{E}}}} a_{ik} a_{jk}, & \text{if } (i, j) \in \bar{\mathcal{E}} \setminus \mathcal{E}. \end{cases}$$

**PROOF.** The proof is analogous to the proof of Theorem 3.3.  $\square$

As a corollary to Proposition 4.1 and Theorem 4.2 we can describe how the normalizing constant changes when removing an edge from a graph with maximal clique size at most 3. Similarly as in the proof of Theorem 3.7, the main difficulty lies in defining a good ordering of the nodes. For simplifying

notation we denote the quotient of the normalizing constants for general  $D$  and the identity matrix by  $\bar{I}_G(\delta, D)$ , i.e.,

$$\bar{I}_G(\delta, D) = \frac{I_G(\delta, D)}{I_G(\delta, \mathbb{I}_p)}.$$

As an example, note that  $\bar{I}_{H_{2,m}}(\delta, D)$  is given in Theorem 4.1.

**COROLLARY 4.3.** *Let  $G = (V, E)$  be an undirected graph with vertices  $V = \{1, \dots, p\}$  and maximal clique size at most 3. Let  $G^e = (V, E^e)$  denote the graph  $G$  with one additional edge  $e$ , i.e.,  $E^e = E \cup \{e\}$ , such that its maximal clique size is also at most 3. Let  $d$  denote the number of triangles formed by the edge  $e$  and two other edges in  $G^e$ . Then*

$$I_G(\delta, D) = \pi^{-1/2} \frac{\Gamma(\delta + \frac{1}{2}(d+2))}{\Gamma(\delta + \frac{1}{2}(d+3))} \frac{|D_{\{1,d+2\}}|^{d-1}}{\prod_{j=2}^{d+1} |D_{\{1,j,d+2\}}|} \bar{I}_{H_{2,d}}(\delta, D) I_{G^e}(\delta, D),$$

where  $D_{\{i_1, \dots, i_k\}}$  denotes the principal submatrix of  $D$  corresponding to the rows and columns  $i_1, \dots, i_k$ .

**PROOF.** We define an ordering of the vertices in such a way that the integral for the normalizing constant with respect to  $G$  decomposes into an integral over a bipartite graph and an integral over the remaining variables. Similarly as in the proof of Theorem 3.7, let one of the end points of  $e$  be labelled as ‘1’, the other end point as ‘ $d+2$ ’ and label the  $d$  vertices involved in triangles over the edge  $e$  by  $2, \dots, d+1$ . Label all remaining vertices by  $d+3, \dots, p$ . Let  $\bar{G}$  denote the moral DAG to  $G$  with edge set  $\bar{E}$  and similarly for  $G^e$ .

By Theorem 4.2, the normalizing constant for  $G$  decomposes into an integral over the variables  $A = \{a_{ij} \mid (i, j) \in \bar{E}, i, j \leq d+2\}$  and an integral over the variables  $B = \{a_{ij} \mid (i, j) \in \bar{E}, a_{ij} \notin A\}$ . The equivalent statement holds for the graph  $G^e$  with  $A^e = A \cup \{e\}$  and  $B^e = B$ . Note that the integral over  $B$  is the same for  $G$  as for  $G^e$ . The integral over  $A$  is the normalizing constant for the complete bipartite graph  $H_{2,d}$  with  $U = \{1, d+2\}$  and  $V = \{2, \dots, d+1\}$  where every vertex in  $U$  is connected to all vertices in  $V$ , but there are no edges within  $U$  nor within  $V$ . The integral over  $A^e = A \cup \{e\}$  is the normalizing constant for the complete bipartite graph  $H_{2,d}$  with one additional edge connecting the two nodes in  $U$ . We denote

this graph by  $H_{2,d}^e$ . So

$$\begin{aligned} I_G(\delta, D) &= I_{G^e}(\delta, D) \frac{I_{H_{2,d}}(\delta, D)}{I_{H_{2,d}^e}(\delta, D)} \\ &= I_{G^e}(\delta, D) \frac{I_{H_{2,d}}(\delta, \mathbb{I}_{d+2}) \bar{I}_{H_{2,d}}(\delta, D)}{I_{H_{2,d}^e}(\delta, D)}, \end{aligned}$$

where  $\bar{I}_{H_{2,d}}(\delta, D)$  is given by Proposition 4.1.

The additional edge  $e$  makes the graph  $H_{2,m}^e$  chordal and hence the normalizing constant can be computed using (1.4):

$$I_{H_{2,d}^e}(\delta, D) = I_{H_{2,d}^e}(\delta, \mathbb{I}_{d+2}) \frac{\prod_{j=2}^{d+1} |D_{\{1,j,d+2\}}|}{|D_{\{1,d+2\}}|^{d-1}}.$$

By Theorem 3.7,

$$\frac{I_{H_{2,d}}(\delta, \mathbb{I}_{d+2})}{I_{H_{2,d}^e}(\delta, \mathbb{I}_{d+2})} = \pi^{-1/2} \frac{\Gamma(\delta + \frac{1}{2}(d+2))}{\Gamma(\delta + \frac{1}{2}(d+3))}.$$

By collecting all terms we find

$$\begin{aligned} I_G(\delta, D) &= I_{G^e}(\delta, D) \frac{I_{H_{2,d}}(\delta, \mathbb{I}_{d+2}) \bar{I}_{H_{2,d}}(\delta, D)}{I_{H_{2,d}^e}(\delta, \mathbb{I}_{d+2})} \frac{|D_{1,d+2}|^{d-1}}{\prod_{j=2}^{d+1} |D_{1,j,d+2}|} \\ &= \pi^{-1/2} \frac{\Gamma(\delta + \frac{1}{2}(d+2))}{\Gamma(\delta + \frac{1}{2}(d+3))} \frac{|D_{\{1,d+2\}}|^{d-1}}{\prod_{j=2}^{d+1} |D_{\{1,j,d+2\}}|} \bar{I}_{H_{2,d}}(\delta, D) I_{G^e}(\delta, D). \end{aligned}$$

The proof now is complete.  $\square$

For graphs with treewidth at most 2, there exists a chordal cover with maximal clique size at most 3. Since by removing edges the maximal clique size cannot increase, Corollary 4.3 can be applied to compute  $I_G(\delta, D)$  for any graph with treewidth at most 2. In order to give a procedure to compute  $I_G(\delta, D)$  for graphs with larger treewidth, we need to generalize Proposition 4.1 to block matrices of the form

$$K = \begin{pmatrix} K_{AA} & K_{AB} \\ K_{AB}^T & K_{BB} \end{pmatrix} \in \mathbb{S}_{\succ 0}^p$$

where  $K_{AA}$  is arbitrary of size  $2 \times 2$ ,  $K_{AB}$  is complete of size  $2 \times m$  and  $K_{BB}$  is arbitrary of size  $m \times m$  and then build on such a result to generalize Corollary 4.3. In Lemma 4.4 we analyze the case in which  $K_{AA}$  is complete and in Lemma 4.5 the case in which  $K_{AA}$  is diagonal. In the following, we denote by  $G_B$  the subgraph of  $G$  induced by the vertices  $B \subset V$ .

LEMMA 4.4. *Let  $G$  be a graph on  $2+m$  vertices with two nodes that are connected to each other and to all other nodes, i.e.  $K$  is of the form*

$$K = \begin{pmatrix} K_{AA} & K_{AB} \\ K_{AB}^T & K_{BB} \end{pmatrix} \in \mathbb{S}_{>0}^{2+m},$$

where  $K_{AA}$  is a complete  $2 \times 2$  matrix,  $K_{AB}$  is a complete  $2 \times m$  matrix and  $K_{BB}$  is an arbitrary  $m \times m$  matrix. Then the integral  $I_G(\delta, D)$  converges absolutely for all  $\delta > -1$  and  $D \in \mathbb{S}_{>0}^{2+m}$ . Further,

$$I_G(\delta, D) = \pi^m \Gamma_2\left(\delta + \frac{3}{2}\right) |D_{AA}|^{-(\delta + \frac{1}{2}(m+3))} \cdot I_{G_B}(\delta + 1, D_{BB} - D_{AB}^T D_{AA}^{-1} D_{AB}).$$

PROOF. By applying the determinant formula for block matrices, making a change-of-variables to replace  $K_{AB}$  by  $K_{AA}^{1/2} K_{AB} K_{BB}^{1/2}$  and applying (4.3) as in the proof of Proposition 4.1 we find that

$$\begin{aligned} I_G(\delta, D) &= \frac{\pi^m \Gamma_2\left(\delta + \frac{3}{2}\right)}{\Gamma_2\left(\delta + \frac{1}{2}(m+3)\right)} \\ &\quad \cdot \int |K_{AA}|^{\delta + \frac{1}{2}m} |K_{BB}|^{\delta+1} \exp(-\text{tr}(K_{AA} D_{AA}) - \text{tr}(K_{BB} D_{BB})) \\ &\quad \cdot {}_0F_1\left(\delta + \frac{1}{2}(m+3); K_{AA} D_{AB} K_{BB} D_{AB}^T\right) dK_{AA} dK_{BB}. \end{aligned}$$

Since  $K_{AA}$  is complete, we can apply (4.2):

$$\begin{aligned} I_G(\delta, D) &= \pi^m \Gamma_2\left(\delta + \frac{3}{2}\right) |D_{AA}|^{-(\delta + \frac{1}{2}(m+3))} \\ &\quad \cdot \int |K_{BB}|^{\delta+1} \exp(-\text{tr}(K_{BB}(D_{BB} - D_{AB}^T D_{AA}^{-1} D_{AB}))) dK_{BB} \\ &= \pi^m \Gamma_2\left(\delta + \frac{3}{2}\right) |D_{AA}|^{-(\delta + \frac{1}{2}(m+3))} \\ &\quad \cdot I_{G_B}(\delta + 1, D_{BB} - D_{AB}^T D_{AA}^{-1} D_{AB}). \end{aligned}$$

This completes the proof.  $\square$

In the following lemma, we will encounter a symmetric matrix  $E_{BB} = (E_{ij})_{i,j \in B}$ . Denoting Kronecker's delta by  $\delta_{ij}$ , we define the matrix of differential operators,

$$\frac{\partial}{\partial E_{BB}} = \left( \frac{1}{2}(1 + \delta_{ij}) \frac{\partial}{\partial E_{ij}} \right)_{i,j \in B}$$

and denote its minor corresponding to the rows  $\{\alpha_1, \alpha_2\}$  and the columns  $\{\beta_1, \beta_2\}$  by

$$\left| \left( \frac{\partial}{\partial E_{BB}} \right)_{\{\alpha_1, \alpha_2\}, \{\beta_1, \beta_2\}} \right|.$$

LEMMA 4.5. *Let  $G$  be a graph on  $2 + m$  vertices with two nodes that are connected to all other nodes but not to each other, i.e.  $K$  is of the form*

$$K = \begin{pmatrix} K_{AA} & K_{AB} \\ K_{AB}^T & K_{BB} \end{pmatrix} \in \mathbb{S}_{>0}^{m+2},$$

where  $K_{AA} = \text{diag}(\kappa_1, \kappa_2)$ ,  $K_{AB}$  is a complete  $2 \times m$  matrix, and  $K_{BB}$  is an arbitrary  $m \times m$  matrix. Then the integral  $I_G(\delta, D)$  converges absolutely for all  $\delta > -1$  and  $D \in \mathbb{S}_{>0}^{2+m}$ , and  $I_G(\delta, D)$  is given by

$$\begin{aligned} & \frac{\pi^m \Gamma_2(\delta + \frac{3}{2})}{\Gamma_2(\delta + \frac{1}{2}(m+3))} [\Gamma(\delta + \frac{1}{2}(m+2))]^2 \sum_{q=0}^{\infty} \frac{(\delta + \frac{1}{2}(m+2))_q}{q! (\delta + \frac{1}{2}(m+3))_{2q}} \\ & \cdot \sum_{l=0}^{\infty} \frac{1}{l! (\delta + 2q + \frac{1}{2}(m+3))_l} \sum_{l_1+l_2=l} \binom{l}{l_1} \left( \prod_{i=1}^2 (\delta + q + \frac{1}{2}(m+2))_{l_i} \right) \\ & \cdot \sum_Q \binom{q}{Q} \left( \prod_{3 \leq \alpha_1 < \alpha_2 \leq m+2} |D_{A, \{\alpha_1, \alpha_2\}}|^{Q_{\alpha_1 \alpha_2}++} \right) \left( \prod_{3 \leq \beta_1 < \beta_2 \leq m+2} |D_{A, \{\beta_1, \beta_2\}}|^{Q_{++\beta_1 \beta_2}} \right) \\ & \cdot \left( \frac{\partial}{\partial t_1} \right)^{l_1} \left( \frac{\partial}{\partial t_2} \right)^{l_2} \left( \prod_{\substack{3 \leq \alpha_1 < \alpha_2 \leq m+2 \\ 3 \leq \beta_1 < \beta_2 \leq m+2}} \left| \left( \frac{\partial}{\partial E_{BB}} \right)_{\{\alpha_1 \alpha_2\}, \{\beta_1 \beta_2\}} \right|^{Q_{\alpha_1 \alpha_2 \beta_1 \beta_2}} \right) \\ & \cdot I_{G_B}(\delta + 1, E_{BB}) \Big|_{E_{BB}=D_{BB}-\sum_{j=1}^2 t_j D_{\{j\}, B}^T D_{\{j\}, B}} \Big|_{t_1=t_2=0}, \end{aligned}$$

where  $Q = (Q_{\alpha_1 \alpha_2 \beta_1 \beta_2} : 3 \leq \alpha_1 < \alpha_2 \leq m+2, 3 \leq \beta_1 < \beta_2 \leq m+2)$  is a vector of non-negative integers such that  $Q_{++++} = q$ .

PROOF. By applying the determinant formula for block matrices, making a change of variables to replace  $K_{AB}$  by  $K_{AA}^{1/2} K_{AB} K_{BB}^{1/2}$  and applying (4.3) as in the proof of Proposition 4.1 we find that

$$\begin{aligned} I_G(\delta, D) &= \frac{\pi^m \Gamma_2(\delta + \frac{3}{2})}{\Gamma_2(\delta + \frac{1}{2}(m+3))} \\ & \cdot \int |K_{AA}|^{\delta + \frac{1}{2}m} |K_{BB}|^{\delta+1} \exp(-\text{tr}(K_{AA} D_{AA}) - \text{tr}(K_{BB} D_{BB})) \\ & \cdot {}_0F_1(\delta + \frac{1}{2}(m+3); K_{AA} D_{AB} K_{BB} D_{AB}^T) dK_{AA} dK_{BB}. \end{aligned}$$

Applying (4.5) and (4.1) as in the proof of Proposition 4.1 we obtain

$$\begin{aligned} & {}_0F_1\left(\delta + \frac{1}{2}(m+3); K_{AA}D_{AB}K_{BB}D_{AB}^T\right) \\ &= \sum_{q=0}^{\infty} \frac{1}{q! \left(\delta + \frac{1}{2}(m+3)\right)_{2q} \left(\delta + \frac{1}{2}(m+2)\right)_q} |K_{AA}D_{AB}K_{BB}D_{AB}^T|^q \\ &\quad \cdot \sum_{l=0}^{\infty} \frac{1}{l! \left(\delta + 2q + \frac{1}{2}(m+3)\right)_l} \left(\text{tr}(K_{AA}D_{AB}K_{BB}D_{AB}^T)\right)^l. \end{aligned}$$

Since  $K_{AA} = \text{diag}(\kappa_1, \kappa_2)$ ,

$$\begin{aligned} \text{tr}(K_{AA}D_{AB}K_{BB}D_{AB}^T) &= \kappa_1 \text{tr}(K_{BB}D_{\{1\},B}^T D_{\{1\},B}) + \kappa_2 \text{tr}(K_{BB}D_{\{2\},B}^T D_{\{2\},B}) \end{aligned}$$

and hence by the Multinomial Theorem

$$\begin{aligned} \left(\text{tr}(K_{AA}K_{AB}K_{BB}K_{AB}^T)\right)^l &= \sum_{l_1+l_2=l} \binom{l}{l_1} \kappa_1^{l_1} \kappa_2^{l_2} \left(\text{tr}(K_{BB}D_{\{1\},B}^T D_{\{1\},B})\right)^{l_1} \\ &\quad \cdot \left(\text{tr}(K_{BB}D_{\{2\},B}^T D_{\{2\},B})\right)^{l_2}. \end{aligned}$$

By the Binet-Cauchy formula ([16, p. 1]),

$$\begin{aligned} & |K_{AA}D_{AB}K_{BB}D_{AB}^T| \\ &= |K_{AA}| \cdot |D_{AB}K_{BB}D_{AB}^T| \\ &= \kappa_1 \kappa_2 \sum_{\substack{3 \leq \alpha_1 < \alpha_2 \leq m+2 \\ 3 \leq \beta_1 < \beta_2 \leq m+2}} |D_{A,\{\alpha_1, \alpha_2\}}| |K_{\{\alpha_1, \alpha_2\}, \{\beta_1, \beta_2\}}| |D_{A,\{\beta_1, \beta_2\}}|. \end{aligned}$$

Hence by the Multinomial Theorem,

$$\begin{aligned} & |K_{AA}D_{AB}K_{BB}D_{AB}^T|^q \\ &= \kappa_1^q \kappa_2^q \sum_Q \binom{q}{Q} \prod_{\substack{3 \leq \alpha_1 < \alpha_2 \leq m+2 \\ 3 \leq \beta_1 < \beta_2 \leq m+2}} |D_{A,\{\alpha_1, \alpha_2\}}|^{Q_{\alpha_1 \alpha_2 \beta_1 \beta_2}} \\ &\quad \cdot |K_{\{\alpha_1, \alpha_2\}, \{\beta_1, \beta_2\}}|^{Q_{\alpha_1 \alpha_2 \beta_1 \beta_2}} |D_{A,\{\beta_1, \beta_2\}}|^{Q_{\alpha_1 \alpha_2 \beta_1 \beta_2}} \\ &= \kappa_1^q \kappa_2^q \sum_Q \binom{q}{Q} \left( \prod_{3 \leq \alpha_1 < \alpha_2 \leq m+2} |D_{A,\{\alpha_1, \alpha_2\}}|^{Q_{\alpha_1 \alpha_2 ++}} \right) \\ &\quad \cdot \left( \prod_{3 \leq \alpha_1 < \alpha_2 \leq m+2} |D_{A,\{\beta_1, \beta_2\}}|^{Q_{++ \beta_1 \beta_2}} \right) \\ &\quad \cdot \left( \prod_{\substack{3 \leq \alpha_1 < \alpha_2 \leq m+2 \\ 3 \leq \beta_1 < \beta_2 \leq m+2}} |K_{\{\alpha_1, \alpha_2\}, \{\beta_1, \beta_2\}}|^{Q_{\alpha_1 \alpha_2 \beta_1 \beta_2}} \right). \end{aligned}$$

Collecting all terms, we find that

$$\begin{aligned}
I_G(\delta, D) &= \frac{\pi^m \Gamma_2(\delta + \frac{3}{2})}{\Gamma_2(\delta + \frac{1}{2}(m+3))} \sum_{q=0}^{\infty} \frac{1}{q! (\delta + \frac{1}{2}(m+3))_{2q} (\delta + \frac{1}{2}(m+2))_q} \\
&\cdot \sum_{l=0}^{\infty} \frac{1}{l! (\delta + 2q + \frac{1}{2}(m+3))_l} \sum_{l_1+l_2=l} \binom{l}{l_1} \left( \prod_{i=1}^2 \int_0^{\infty} \kappa_i^{\delta+q+l_i+\frac{1}{2}m} e^{-\kappa_i} d\kappa_i \right) \\
&\cdot \sum_Q \binom{q}{Q} \left( \prod_{3 \leq \alpha_1 < \alpha_2 \leq m+2} |D_{A, \{\alpha_1, \alpha_2\}}|^{Q_{\alpha_1 \alpha_2 ++}} \right) \\
&\cdot \left( \prod_{3 \leq \beta_1 < \beta_2 \leq m+2} |D_{A, \{\beta_1, \beta_2\}}|^{Q_{++ \beta_1 \beta_2}} \right) \\
&\cdot \int |K_{BB}|^{\delta+1} \exp(-\text{tr}(K_{BB} D_{BB})) \left( \prod_{j=1}^2 (\text{tr}(K_{BB} D_{\{j\}, B}^T D_{\{j\}, B}))^{l_j} \right) \\
&\cdot \left( \prod_{\substack{3 \leq \alpha_1 < \alpha_2 \leq m+2 \\ 3 \leq \beta_1 < \beta_2 \leq m+2}} |K_{\{\alpha_1, \alpha_2\}, \{\beta_1, \beta_2\}}|^{Q_{\alpha_1 \alpha_2 \beta_1 \beta_2}} \right) dK_{BB}.
\end{aligned}$$

The two gamma integrals are computed easily, so only the integral over the variables  $K_{BB}$  remains to be evaluated, and we shall evaluate that integral in terms of a normalizing constant for the graph  $G_B$ . First, note that

$$\begin{aligned}
&\exp(-\text{tr}(K_{BB} D_{BB})) \prod_{j=1}^2 \left( \text{tr}(K_{BB} D_{\{j\}, B}^T D_{\{j\}, B}) \right)^{l_j} \\
&= \left( \frac{\partial}{\partial t_1} \right)^{l_1} \left( \frac{\partial}{\partial t_2} \right)^{l_2} \exp[-\text{tr}(K_{BB} E_{BB})] \Big|_{t_1=t_2=0},
\end{aligned}$$

where

$$E_{BB} = D_{BB} - \sum_{j=1}^2 t_j D_{\{j\}, B}^T D_{\{j\}, B}.$$

Let  $E_{ij}$  denote the entry of  $E_{BB}$  corresponding to nodes  $i$  and  $j$  in  $B$ . Then

$$\begin{aligned}
&\int |K_{BB}|^{\delta+1} \exp(-\text{tr}(K_{BB} E_{BB})) |K_{\{\alpha_1, \alpha_2\}, \{\beta_1, \beta_2\}}|^{Q_{\alpha_1 \alpha_2 \beta_1 \beta_2}} dK_{BB} \\
&= \left| \left( \frac{\partial}{\partial E_{BB}} \right)_{\{\alpha_1 \alpha_2\}, \{\beta_1 \beta_2\}} \right|^{Q_{\alpha_1 \alpha_2 \beta_1 \beta_2}} \int |K_{BB}|^{\delta+1} \exp(-\text{tr}(K_{BB} E_{BB})) dK_{BB}.
\end{aligned}$$

By collecting all terms, we obtain the desired result.  $\square$

With these two lemmas, we now have the tools to generalize Corollary 4.3 to graphs of treewidth larger than 2. In the following theorem, we will show how the normalizing constant changes when removing an edge from a general graph  $G$ .

**THEOREM 4.6.** *Let  $G = (V, E)$  be an undirected graph on  $p$  vertices. Let  $G^e = (V, E^e)$  denote the graph  $G$  with one additional edge  $e$ , i.e.,  $E^e = E \cup \{e\}$ . Let  $d$  denote the number of triangles formed by the edge  $e$  and two other edges in  $G^e$ . Let  $V$  be partitioned such that  $V = A \cup B \cup C$  with  $|A| = 2$ ,  $|B| = d$ ,  $|C| = p - d - 2$ , and where  $A$  contains the two vertices adjacent to the edge  $e$  in  $G^e$ ,  $B$  contains all vertices in  $G^e$  that span a triangle with the edge  $e$ , and  $C$  contains all remaining nodes. Then  $I_G(\delta, D)$  is given by*

$$\begin{aligned} & \pi^{-1/2} \frac{\Gamma(\delta + \frac{1}{2}(d+2))}{\Gamma(\delta + \frac{1}{2}(d+3))} I_{G^e}(\delta, D) |D_{AA}|^{\delta + \frac{1}{2}(d+3)} \sum_{q=0}^{\infty} \frac{(\delta + \frac{1}{2}(d+2))_q}{q! (\delta + \frac{1}{2}(d+3))_{2q}} \\ & \cdot \sum_{l=0}^{\infty} \frac{1}{l! (\delta + 2q + \frac{1}{2}(d+3))_l} \sum_{l_1+l_2=l} \binom{l}{l_1} \prod_{i=1}^2 (\delta + q + \frac{1}{2}(d+2))_{l_i} \\ & \cdot \sum_Q \binom{q}{Q} \left( \prod_{3 \leq \alpha_1 < \alpha_2 \leq d+2} |D_{A, \{\alpha_1, \alpha_2\}}|^{Q_{\alpha_1 \alpha_2 ++}} \right) \left( \prod_{3 \leq \beta_1 < \beta_2 \leq d+2} |D_{A, \{\beta_1, \beta_2\}}|^{Q_{++ \beta_1 \beta_2}} \right) \\ & \cdot \left( \frac{\partial}{\partial t_1} \right)^{l_1} \left( \frac{\partial}{\partial t_2} \right)^{l_2} \left( \prod_{\substack{3 \leq \alpha_1 < \alpha_2 \leq d+2 \\ 3 \leq \beta_1 < \beta_2 \leq d+2}} \left| \left( \frac{\partial}{\partial E_{BB}} \right)_{\{\alpha_1 \alpha_2\}, \{\beta_1 \beta_2\}} \right|^{Q_{\alpha_1 \alpha_2 \beta_1 \beta_2}} \right) \\ & \cdot \frac{I_{G_B}(\delta + 1, E_{BB})}{I_{G_B}(\delta + 1, D_{BB} - D_{AB}^T D_{AA}^{-1} D_{AB})} \Big|_{E_{BB}=D_{BB}-\sum_{j=1}^2 t_j D_{\{j\}, B}^T D_{\{j\}, B}} \Big|_{t_1=t_2=0}. \end{aligned}$$

**PROOF.** Let  $G_{AB}$  be the graph induced by the vertices  $A \cup B$ . By Theorem 4.2 and as in the proof of Corollary 4.3, the normalizing constants for  $G$  and  $G^e$  decompose into the normalizing constants for  $G_{AB}$  and  $G_{AB}^e$ , respectively, and an integral over the variables involving  $C$ . Moreover, the integral over the variables involving  $C$  is the same for  $G$  and for  $G^e$ . Hence,

$$\frac{I_G(\delta, D)}{I_{G^e}(\delta, D)} = \frac{I_{G_{AB}}(\delta, D_{\{A, B\}, \{A, B\}})}{I_{G_{AB}^e}(\delta, D_{\{A, B\}, \{A, B\}})},$$

where  $D_{\{A, B\}, \{A, B\}}$  denotes the principle submatrix of  $D$  corresponding to the rows and columns in  $A \cup B$ . Since  $G_{AB}$  is of the form needed for Lemma 4.5 and  $G_{AB}^e$  is of the form needed for Lemma 4.4, the claim follows by applying Lemma 4.4 and Lemma 4.5.  $\square$



For the case in which  $D$  is the identity matrix,  $E_{BB} \equiv D_{BB}$  and hence for  $l_1 > 0$  or  $l_2 > 0$  we get

$$\left(\frac{\partial}{\partial t_1}\right)^{l_1} \left(\frac{\partial}{\partial t_2}\right)^{l_2} \left( \prod_{\substack{3 \leq \alpha_1 < \alpha_2 \leq d+2 \\ 3 \leq \beta_1 < \beta_2 \leq d+2}} \left| \left(\frac{\partial}{\partial E_{BB}}\right)_{\{\alpha_1 \alpha_2\}, \{\beta_1 \beta_2\}} \right|^{Q_{\alpha_1 \alpha_2 \beta_1 \beta_2}} \right) \cdot \frac{I_{G_B}(\delta + 1, E_{BB})}{I_{G_B}(\delta + 1, D_{BB} - D_{AB}^T D_{AA}^{-1} D_{AB})} = 0.$$

In addition,  $|D_{A, \{\alpha_1, \alpha_2\}}| = |D_{A, \{\beta_1, \beta_2\}}| = 0$ . Hence, in the infinite sums only the terms for  $q = 0$  and  $l = 0$  are non-zero, and the infinite series reduce to 1. Since  $|D_{AA}| = 1$ , we see that if  $D = \mathbb{I}_p$  then Theorem 4.6 reduces to Theorem 3.7.

By applying this theorem, we can obtain an exact formula for the normalizing constant  $I_G(\delta, D)$  for any graph  $G$ . First, note that if  $G^e$  has treewidth  $m$ , then the induced graph  $G_B$ , where  $B \subset V$  contains all vertices in  $G^e$  that span a triangle with the edge  $e$ , has treewidth at most  $m - 2$ . Corollary 4.3 leads to a procedure that allows computing the normalizing constant for any graph of treewidth at most 2. Hence by applying Theorem 4.6, we can compute the normalizing constant for any graph of treewidth at most 4. By repeatedly applying Theorem 4.6 we can compute the normalizing constant for any graph.

More precisely, the procedure is the following: Given a graph  $G$ , it is desirable to compute a minimal chordal cover of  $G$ . Here, minimality means that we choose a chordal cover with minimal clique size and among all these chordal covers we choose one with the minimal number of added edges. Then we delete, one at a time, the added edges that are in the minimal chordal cover of  $G$  but not in  $G$  itself, and update the normalizing constant at every step. For each such edge, updating the normalizing constant requires computing the normalizing constant for the induced graph  $G_B$ , where  $B \subset V$  contains all vertices in  $G^e$  that span a triangle with the edge  $e$ . So we reduced the problem to a simpler problem since

$$\text{treewidth}(G_B) \leq \text{treewidth}(G) - 2.$$

If the treewidth of  $G_B$  is at most 2, then applying Corollary 4.3 gives the normalizing constant for  $G_B$ . Otherwise, we need to compute the normalizing constant for  $G_B$  by taking a minimal chordal cover of  $G_B$  and removing one edge at a time.

This procedure allows computing the normalizing constant for any graph. However, for graphs of large treewidth this procedure may be quite expensive

from a computational point of view. In the following example, we show how this procedure can be applied to compute the normalizing constant of  $G_5$ .

EXAMPLE 4.7. We revisit the example  $G_5$  discussed in Section 2 and Example 3.8. We will obtain the normalizing constant,  $I_{G_5}(\delta, D)$ , explicitly and show that it reduces for the case in which  $D = \mathbb{I}_5$  to the result given in Example 3.8.

A minimal chordal cover of  $G_5$  is given in Figure 1 (right). Only one edge is in the chordal cover of  $G_5$  but not in  $G_5$  itself, namely the edge  $e = (1, 3)$ . We denote the chordal cover of  $G_5$  by  $G_5^e$ . There are  $d = 3$  triangles formed by the edge  $e$  in  $G_5^e$ . The vertices adjacent to the edge  $e$  are  $A = \{1, 3\}$  and all remaining vertices span a triangle with the edge  $e$ , i.e.  $B = \{2, 4, 5\}$ . The induced graph  $G_B$  consists of one edge only, namely  $(4, 5)$ , so its normalizing constant is

$$I_{G_B}(\delta, D_B) = |D_{\{4,5\}}|^{-(\delta+\frac{3}{2})} \Gamma_2(\delta + \frac{3}{2}) \Gamma(\delta + 1),$$

where, in order to abbreviate notation, we denoted by  $D_B$  and  $D_{\{4,5\}}$ , respectively, the principle submatrix of  $D$  corresponding to the rows and columns in  $B$  and in  $\{4, 5\}$ , respectively. Hence by applying Theorem 4.6, we obtain the following formula for  $I_{G_5}(\delta, D)$ :

$$\begin{aligned} & \pi^{-1/2} \frac{\Gamma(\delta + \frac{5}{2})}{\Gamma(\delta + 3)} I_{G_5^e}(\delta, D) |D_{AA}|^{\delta+3} \sum_{q=0}^{\infty} \frac{(\delta + \frac{5}{2})_q}{q! (\delta + 3)_{2q}} \\ & \cdot \sum_{l=0}^{\infty} \frac{1}{l! (\delta + 2q + 3)_l} \sum_{l_1+l_2=l} \binom{l}{l_1} \prod_{i=1}^2 (\delta + q + \frac{5}{2})_{l_i} \\ & \cdot \sum_Q \binom{q}{Q} \left( \prod_{\{\alpha_1, \alpha_2\} \subset \{2,4,5\}} |D_{A, \{\alpha_1, \alpha_2\}}|^{Q_{\alpha_1 \alpha_2}++} \right) \left( \prod_{\{\beta_1, \beta_2\} \subset \{2,4,5\}} |D_{A, \{\beta_1, \beta_2\}}|^{Q_{++\beta_1 \beta_2}} \right) \\ & \cdot \left( \frac{\partial}{\partial t_1} \right)^{l_1} \left( \frac{\partial}{\partial t_2} \right)^{l_2} \left( \prod_{\substack{\{\alpha_1, \alpha_2\} \subset \{2,4,5\} \\ \{\beta_1, \beta_2\} \subset \{2,4,5\}}} \left| \left( \frac{\partial}{\partial E_{BB}} \right)_{\{\alpha_1 \alpha_2\}, \{\beta_1 \beta_2\}} \right|^{Q_{\alpha_1 \alpha_2 \beta_1 \beta_2}} \right) \\ & \cdot \frac{|D_{\{4,5\}} - D_{A, \{4,5\}}^T D_{AA}^{-1} D_{A, \{4,5\}}|^{\delta+\frac{3}{2}}}{|E_{\{4,5\}}|^{\delta+\frac{3}{2}}} \Bigg|_{E_{BB}=D_{BB}-\sum_{j=1}^2 t_j D_{\{j\},B}^T D_{\{j\},B}} \Bigg|_{t_1=t_2=0}. \end{aligned}$$

Note that when  $2 \in \{\alpha_1, \alpha_2\}$  or  $2 \in \{\beta_1, \beta_2\}$  and  $Q_{\alpha_1 \alpha_2 \beta_1 \beta_2} > 0$ , then

$$\left| \left( \frac{\partial}{\partial E_{BB}} \right)_{\{\alpha_1 \alpha_2\}, \{\beta_1 \beta_2\}} \right|^{Q_{\alpha_1 \alpha_2 \beta_1 \beta_2}} |E_{\{4,5\}}|^{-(\delta+\frac{3}{2})} = 0.$$

As a consequence, the normalizing constant for  $I_{G_5}(\delta, D)$  is given by

$$\begin{aligned} & \pi^{-1/2} \frac{\Gamma(\delta + \frac{5}{2})}{\Gamma(\delta + 3)} I_{G_5^e}(\delta, D) |D_{AA}|^{\delta+3} \sum_{q=0}^{\infty} \frac{(\delta + \frac{5}{2})_q}{q! (\delta + 3)_{2q}} |D_{A,\{4,5\}}|^{2q} \\ & \cdot \sum_{l=0}^{\infty} \frac{1}{l! (\delta + 2q + 3)_l} \sum_{l_1+l_2=l} \binom{l}{l_1} \prod_{i=1}^2 (\delta + q + \frac{5}{2})_{l_i} \\ & \cdot \left( \frac{\partial}{\partial t_1} \right)^{l_1} \left( \frac{\partial}{\partial t_2} \right)^{l_2} \left| \frac{\partial}{\partial E_{\{4,5\}}} \right|^q \\ & \cdot \frac{|D_{\{4,5\}} - D_{A,\{4,5\}}^T D_{AA}^{-1} D_{A,\{4,5\}}|^{\delta+\frac{3}{2}}}{|E_{\{4,5\}}|^{\delta+\frac{3}{2}}} \bigg|_{\substack{E_{\{4,5\}}=D_{\{4,5\}}-\sum_{j=1}^2 t_j D_{\{j\},\{4,5\}}^T D_{\{j\},\{4,5\}} \\ t_1=t_2=0}}, \end{aligned}$$

the evaluation being done first at  $E_{\{4,5\}} = D_{\{4,5\}} - \sum_{j=1}^2 t_j D_{\{j\},\{4,5\}}^T D_{\{j\},\{4,5\}}$  and last at  $t_1 = t_2 = 0$ . Since  $G_5^e$  is chordal, the corresponding normalizing constant is obtained from (1.4):

$$I_{G_5^e}(\delta, D) = \Gamma_3(\delta + 2) \Gamma_4(\delta + \frac{5}{2}) \frac{|D_{\{1,2,3\}}|^{-(\delta+2)} |D_{\{1,3,4,5\}}|^{-(\delta+\frac{5}{2})}}{|D_{\{1,3\}}|^{-(\delta+\frac{3}{2})} \Gamma_2(\delta + \frac{3}{2})}.$$

Note also that

$$\left| \frac{\partial}{\partial E_{\{4,5\}}} \right|^q |E_{\{4,5\}}|^{-(\delta+\frac{3}{2})} = (-1)^{2q} \frac{\Gamma_2(\delta + \frac{3}{2} + q)}{\Gamma_2(\delta + \frac{3}{2})} |E_{\{4,5\}}|^{-(\delta+\frac{3}{2}+q)};$$

this can be obtained by writing  $I_{\text{complete}}(\delta, D)$  as an integral in Equation (1.2) and applying the differential operator  $\partial/\partial E_{\{4,5\}}$  to both sides of the equation (Maass [21], p. 81). Hence, by collecting all terms, noting that  $A = \{1, 3\}$  and simplifying the formula for  $I_{G_5}(\delta, D)$  above, we obtain the normalizing constant for  $G_5$  for general  $D$ :

$$\begin{aligned} I_{G_5}(\delta, D) &= I_{G_5}(\delta, \mathbb{I}_5) \\ & \cdot \frac{|D_{\{1,3\}}|^{2\delta+\frac{9}{2}}}{|D_{\{1,2,3\}}|^{\delta+2} |D_{\{1,3,4,5\}}|^{\delta+\frac{5}{2}}} \sum_{q=0}^{\infty} \frac{(\delta + \frac{5}{2})_q (\delta + \frac{3}{2})_q}{q! (\delta + 3)_{2q}} \\ & \cdot |D_{\{1,3\}\{4,5\}}|^{2q} \sum_{l=0}^{\infty} \frac{1}{l! (\delta + 2q + 3)_l} \sum_{l_1+l_2=l} \binom{l}{l_1} \prod_{i=1}^2 (\delta + q + \frac{5}{2})_{l_i} \\ & \cdot \left( \frac{\partial}{\partial t_1} \right)^{l_1} \left( \frac{\partial}{\partial t_2} \right)^{l_2} \frac{|D_{\{4,5\}} - D_{\{1,3\},\{4,5\}}^T D_{\{1,3\}}^{-1} D_{\{1,3\},\{4,5\}}|^{\delta+\frac{3}{2}}}{|D_{\{4,5\}} - \sum_{j=1}^2 t_j D_{\{j\},\{4,5\}}^T D_{\{j\},\{4,5\}}|^{\delta+q+\frac{3}{2}}} \bigg|_{t_1=t_2=0}. \end{aligned}$$

If  $D = \mathbb{I}_5$  then  $D_{\{j\},\{4,5\}}^T D_{\{j\},\{4,5\}} = 0$  and hence for  $l_1 > 0$  or  $l_2 > 0$  we get

$$\left(\frac{\partial}{\partial t_1}\right)^{l_1} \left(\frac{\partial}{\partial t_2}\right)^{l_2} \frac{|D_{\{4,5\}} - D_{\{1,3\},\{4,5\}}^T D_{\{1,3\}}^{-1} D_{\{1,3\},\{4,5\}}|^{\delta + \frac{3}{2}}}{|D_{\{4,5\}} - \sum_{j=1}^2 t_j D_{\{j\},\{4,5\}}^T D_{\{j\},\{4,5\}}|^{\delta + q + \frac{3}{2}}} = 0.$$

In addition,  $|D_{\{1,3\},\{4,5\}}| = 0$ . Hence in the infinite sums only the terms for  $q = 0$  and  $l = 0$  are non-zero, and the infinite sums reduce to 1. Since  $|D_{\{1,3\}}| = |D_{\{1,2,3\}}| = |D_{\{1,3,4,5\}}| = 1$ , we see that the formula for  $I_{G_5}(\delta, D)$  indeed reduces to  $I_{G_5}(\delta, \mathbb{I}_5)$ .

### ACKNOWLEDGMENTS

A.L.'s research was supported by Statistics for Innovation *sfi*<sup>2</sup> in Oslo. D.R.'s research was partially supported by the U.S. National Science Foundation grant DMS-1309808; and by a Romberg Guest Professorship at the Heidelberg University Graduate School for Mathematical and Computational Methods in the Sciences, funded by German Universities Excellence Initiative grant GSC 220/2.

### REFERENCES

- [1] ANDERSON, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*, Third ed. Wiley, Hoboken, NJ. [MR1990662](#)
- [2] ANDREWS, G. E., ASKEY, R. and ROY, R. (2000). *Special Functions*. Cambridge University Press, New York. [MR1688958](#)
- [3] ATAY-KAYIS, A. and MASSAM, H. (2005). A Monte Carlo method for computing the marginal likelihood in nondecomposable Gaussian graphical models. *Biometrika* **92** 317–335. [MR2201362](#)
- [4] CHENG, Y. and LENKOSKI, A. (2012). Hierarchical Gaussian graphical models: Beyond reversible jump. *Electronic Journal of Statistics* **6** 2309–2331.
- [5] DAWID, A. P. and LAURITZEN, S. L. (1993). Hyper Markov laws in the statistical analysis of decomposable graphical models. *Ann. Statist.* **21** 1272–1317.
- [6] DEMPSTER, A. P. (1972). Covariance selection. *Biometrics* **28** 157–175.
- [7] DOBRA, A. and LENKOSKI, A. (2011). Copula Gaussian graphical models and their application to modeling functional disability data. *Annals of Applied Statistics* **5** 969–993.
- [8] DOBRA, A., LENKOSKI, A. and RODRIGUEZ, A. (2011). Bayesian inference for general Gaussian graphical models with application to multivariate lattice data. *Journal of the American Statistical Association* **106** 1418–1433.
- [9] GIRI, N. C. (2004). *Multivariate Statistical Analysis*. Marcel Dekker, New York. [MR0468025](#)
- [10] GIUDICI, P. and GREEN, P. J. (1999). Decomposable graphical Gaussian model determination. *Biometrika* **86** 785–801.
- [11] GROSS, K. I. and RICHARDS, D. ST. P. (1987). Special functions of matrix argument. I. Algebraic induction, zonal polynomials, and hypergeometric functions. *Trans. Amer. Math. Soc.* **301** 781–811. [MR0882715](#)

- [12] HERZ, C. S. (1955). Bessel functions of matrix argument. *Ann. Math.* **61** 474–523. [MR0069960](#)
- [13] INGHAM, A. E. (1933). An integral which occurs in statistics. *Math. Proc. Camb. Phil. Soc.* **29** 271–276.
- [14] JAMES, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.* **35** 475–501. [MR0181057](#)
- [15] JONES, B., CARVALHO, C., DOBRA, A., HANS, C., CARTER, C. and WEST, M. (2005). Experiments in stochastic computation for high-dimensional graphical models. *Statistical Science* **20** 388–400.
- [16] KARLIN, S. (1968). *Total Positivity* **1**. Stanford University Press, Stanford, CA. [MR0230102](#)
- [17] LAURITZEN, S. L. (1996). *Graphical Models*. Oxford University Press, New York. [MR1419991](#)
- [18] LENKOSKI, A. (2013). A direct sampler for  $G$ -Wishart variates. *Stat* **2** 119–128.
- [19] LENKOSKI, A. and DOBRA, A. (2011). Computational aspects related to inference in Gaussian graphical models with the  $G$ -Wishart prior. *J. Comput. Graph. Statist.* **20** 140–157. [MR2816542](#)
- [20] LETAC, G. and MASSAM, H. (2007). Wishart distributions for decomposable graphs. *Ann. Statist.* **35** 1278–323.
- [21] MAASS, H. (1971). *Siegel's Modular Forms and Dirichlet Series. Lecture Notes in Mathematics* **216**. Springer, Heidelberg.
- [22] MITSAKAKIS, N., MASSAM, H. and ESCOBAR, M. D. (2011). A Metropolis-Hastings based method for sampling from the  $G$ -Wishart distribution in Gaussian graphical models. *Electronic Journal of Statistics* **5** 18–30.
- [23] MUIRHEAD, R. J. (1975). Expressions for some hypergeometric functions of matrix argument with applications. *J. Multivariate Anal.* **5** 283–293. [MR0381137](#)
- [24] MUIRHEAD, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, Hoboken, NJ. [MR0652932](#)
- [25] OLKIN, I. (2002). The 70th anniversary of the distribution of random matrices: A survey. *Linear Algebra Appl.* **354** 231–243. [MR1927659](#)
- [26] PICCIONI, M. (2000). Independence structure of natural conjugate densities to exponential families and the Gibbs sampler. *Scand. J. Statist.* **27** 111–27.
- [27] ROVERATO, A. (2002). Hyper inverse Wishart distribution for non-decomposable graphs and its application to Bayesian inference for Gaussian graphical models. *Scand. J. Statist.* **29** 391–411. [MR1925566](#)
- [28] SIEGEL, C. L. (1935). Über die analytische Theorie der quadratischen Formen. *Ann. Math.* **36** 527–606. [MR1503238](#)
- [29] SPEED, T. P. and KIIVERI, H. (1986). Gaussian Markov distributions over finite graphs. *Annals of Statistics* **14** 138–150.
- [30] WANG, H. and CARVALHO, C. M. (2010). Simulation of hyper-inverse Wishart distributions for non-decomposable graphs. *Electronic Journal of Statistics* **4** 1470–1475.
- [31] WANG, H. and LI, S. Z. (2012). Efficient Gaussian graphical model determination under  $G$ -Wishart prior distributions. *Electronic Journal of Statistics* **6** 168–198.
- [32] WISHART, J. (1928). The generalised product moment distribution in samples from a normal multivariate population. *Biometrika* **20A** 32–52.
- [33] WISHART, J. and BARTLETT, M. S. (1933). The generalised product moment distribution in a normal system. *Math. Proc. Camb. Phil. Soc.* **29** 260–270.

C. UHLER  
IST ASTRIA  
3400 KLOSTERNEUBURG  
AUSTRIA  
E-MAIL: [caroline.uhler@ist.ac.at](mailto:caroline.uhler@ist.ac.at)

A. LENKOSKI  
NORWEGIAN COMPUTING CENTER  
OSLO, NORWAY  
E-MAIL: [alex@nr.no](mailto:alex@nr.no)

D. RICHARDS  
DEPARTMENT OF STATISTICS  
PENN STATE UNIVERSITY  
UNIVERSITY PARK, PA 16802, U.S.A.  
AND  
INSTITUTE OF APPLIED MATHEMATICS  
HEIDELBERG UNIVERSITY  
INF, HEIDELBERG, GERMANY  
E-MAIL: [richards@stat.psu.edu](mailto:richards@stat.psu.edu)